

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL



**ADESÃO À TERAPÊUTICA ANTI-RETROVÍRICA – VIH:
DETERMINAÇÃO DOS PRINCIPAIS FATORES ASSOCIADOS À
NÃO-ADESÃO À TERAPÊUTICA E CARACTERIZAÇÃO DOS
DIFERENTES PERFIS DE DOENTES**

João Carlos Proença Marques

Trabalho de Projeto
Mestrado em Estatística e Investigação Operacional
Especialização em Estatística

2013

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL



**ADESÃO À TERAPÊUTICA ANTI-RETROVÍRICA – VIH:
DETERMINAÇÃO DOS PRINCIPAIS FATORES ASSOCIADOS À
NÃO-ADESÃO À TERAPÊUTICA E CARACTERIZAÇÃO DOS
DIFERENTES PERFIS DE DOENTES**

João Carlos Proença Marques

Trabalho de Projeto
Mestrado em Estatística e Investigação Operacional
Especialização em Estatística

Orientadoras:

Prof.^a Doutora Margarida Maria Teixeira Diniz Mendes Leal
Mestre Milene Catarina Coelho Fernandes

2013

Agradecimentos

A elaboração desta tese apenas foi possível devido ao apoio de professores, colegas, amigos e familiares. Assim, uma vez que não é possível agradecer de forma devida a todos aqueles que me acompanharam neste percurso, destaco aqueles que considero fundamentais e que deram o seu contributo para o resultado final desta tese.

Agradeço à Professora Doutora Margarida Mendes Leal por me ter orientado e pela disponibilidade que mostrou sempre que, por alguma razão, necessitei do seu apoio.

Gostaria, igualmente, de agradecer à equipa que integra o Instituto de Medicina Preventiva, nomeadamente, à Mestre Milene Fernandes, por me ter orientado em todas as fases deste projeto, dando ideias, propondo desafios, fazendo-me dar sempre um pouco mais de mim. Fez-me sentir parte integrante deste projeto, possibilitando que eu apresentasse ideias e sugestões.

À minha família, nomeadamente a minha mãe e o meu irmão, agradeço todo o apoio que me deram, nos momentos em que, por alguma razão, estava desmotivado, frustrado.

Agradeço a todos os colegas de mestrado que partilharam das minhas dúvidas e angústias sobre a elaboração de uma tese, partilhando ideias e dando sugestões. Em especial agradeço ao Pedro Rosário, por me ter dado ânimo para continuar nos momentos em que julgava não ser capaz. O facto de nos encontrarmos em situações semelhantes fez com que o espírito de entre-ajuda imperasse, permitindo o apoio mútuo.

À Tânia Costa (a minha alma gémea) por ter percorrido este longo caminho comigo desde a entrada na faculdade. Construámos uma enorme amizade que nos ajudou a encarar as adversidades com relativa facilidade. Quando reparava que estava frustrado por algo não resultar como pretendia, preocupou-se em distrair-me, permitindo-me voltar a enfrentar este desafio com um um ânimo diferente.

A todos,o meu muito obrigado.

Resumo

Introdução: O tratamento anti-retrovírico tem contribuído significativamente para a redução da morbidade e mortalidade em indivíduos infetados pelo VIH, principalmente, a partir de 1996, com a introdução da terapia anti-retroviral combinada (TARc). Todavia, o sucesso do regime terapêutico está fortemente dependente de uma adesão (preferencialmente) ótima. A maioria dos estudos efetuados até à data, aceita como ponto ótimo de corte, o uso correto de, pelo menos, 95% das doses prescritas. São vários os métodos e recursos já utilizados para avaliar a adesão à terapêutica anti-retrovírica (TAR).

Objetivos: O presente estudo procura estimar a prevalência da não-adesão à terapêutica anti-retrovírica, através de questionários efetuados aos doentes e relacionar a não-adesão com características socio-demográficas, biomédicas e psicossociais em adultos infetados com VIH-1. Consequentemente, procura-se caracterizar os diferentes perfis de doentes.

Métodos: Foi conduzido um estudo prospetivo, através de uma amostra sistemática de adultos infetados com VIH-1 aos quais foi prescrita a TARc e seguidos no Hospital de Santa Maria, com, pelo menos, uma consulta médica no período de 3 meses (maio-julho 2011). Os doentes foram, posteriormente, entrevistados numa outra consulta médica, num período até 6 meses. A adesão à terapêutica foi avaliada através do questionário *Adults AIDS Clinical Trials Group Adherence questionnaire (AACTG)*. Os fatores relacionados com o doente foram avaliados através de um questionário específico, o qual inclui o Questionário da Perceção da Doença (*IPQ-R*), o Questionário de Crenças sobre a medicação e as escalas de Satisfação com o Suporte Social e a de Depressão, Ansiedade e Stress. Os doentes foram classificados como não aderentes se apresentavam, pelo menos, uma falha na toma da medicação no último fim de semana, nos 4 dias anteriores do questionário ou nos últimos 30 dias que antecederam o mesmo. Foi utilizado um modelo de regressão logística para identificar os fatores associados à não-adesão. O nível de significância utilizado foi de $\alpha=0,05$. Posteriormente, foi utilizada a técnica multivariada de Análise de Correspondências Múltiplas com o mesmo objetivo.

Resultados: De 306 doentes que se apresentaram na consulta do período de 3 meses após o início da terapêutica, 198 (64,7%) mostraram-se elegíveis no consulta inicial,

respondendo ao questionário na totalidade e, de entre estes, 177 completaram o questionário na consulta de seguimento. Os participantes apresentaram uma idade de $48,4 \pm 10,1$ anos, 76,3% homens e encontravam-se na terapêutica à $122,1 \pm 63,95$ meses. 86 doentes (48,6%) foram classificados como não aderentes. Foram incluídos no modelo de regressão logística 177 participantes. O nível de ansiedade, o facto do indivíduos morar sozinho, o controlo do tratamento, a situação clínica no momento da entrevista, as preocupações em relação aos fármacos e o facto do participante ser *binge drinker* mostraram-se independentemente associados à não-adesão ($p \leq 0,05$).

Conclusões: A não adesão à TARc é altamente prevalente nesta amostra de adultos infetados com VIH-1. Foram identificadas variáveis relacionadas com a não-adesão, pelo que deverão ser alvo de controlo de forma a prevenir o desenvolvimento da resistência a terapêutica.

Palavras-chave: Infecção VIH/SIDA, Regressão Logística, Análise de Correspondências Múltiplas, Adesão à Terapêutica Anti-Retroviral.

Abstract

Introduction: Antiretroviral therapy has contributed significantly to the reduction of morbidity and mortality in patients infected with HIV, mainly from 1996, with the introduction of combination antiretroviral therapy (cART). However, the success of the treatment regimen is strongly dependent of optimal adherence. Most studies conducted to date, accepted as optimum cohort, the correct use of, at least, 95% of the prescribed doses. There are several methods and resources already used to assess adherence to antiretroviral therapy (ART).

Aims: This study aimed to estimate the prevalence of patient non-adherence to cART and to identify associated patient-related factors socio-demographic, biomedical and psychosocial in HIV-1 infected Portuguese adults. Consequently, we seek to characterize the different patient profiles.

Methods: A prospective cohort study was conducted, with systematic sampling of HIV-1 infected adults prescribed to cART and followed at the Infectious Disease Outpatient Clinic (Hospital de Santa Maria, Lisbon), with at least one medical appointment during the 3-month enrolment period (May-July 2011). Patients were followed-up until the next medical appointment, for a 6-month period. Adherence was assessed through the Adults AIDS Clinical Trials Group Adherence questionnaire. Patient-related factors were assessed through a specific questionnaire that included the Illness Perception Questionnaire (*IPQ-R*), the Satisfaction with Social Support Scale, Beliefs about Medicines Questionnaire (*BMQ*) and the Anxiety Depression and Stress Scale. At baseline, subjects were classified as non-adherents if they had at least one failure during the last weekend, the 4 days prior to the questionnaire or the 30 days before the questionnaire. Logistic regression model was used to identify factors independently associated with non-adherence. Statistical significance was set at $\alpha=0.05$. Subsequently, we used the multivariate technique of Multiple Correspondence Analysis with the same goal.

Results: From 306 subjects who had a medical appointment during the enrolment period, 198 (64.7%) were eligible at baseline and fullfilled the adherence questionnaire, and 177 completed follow-up. The participants were $48,4 \pm 10,0$ years-old (mean \pm SD),

76,3 % man, and prescribed to cART for $122,1 \pm 63,95$ months (mean \pm SD). A total of 86 (48,6%) patients were non-adherent. Data from 177 patients were included in the regression model. Anxiety, living alone, treatment control, clinical situation at the time of interview, IPQ concerns and *binge-drinking* were independently associated to non-adherence ($p \leq 0,05$).

Conclusions: Non-adherence to cART is highly prevalent in this Portuguese sample of HIV-1 infected adults. Patient-related variables were independently associated to non-adherence and should be addressed, in order to improve cART utilization and to prevent drug resistance development.

Key-words: HIV/AIDS infection, multivariate logistic regression analysis, multiple correspondence analysis, antiretroviral therapy.

Conteúdo

1	Introdução	3
1.1	A infecção VIH/SIDA	3
1.1.1	Terapêutica Anti-Retrovítica (TAR)	6
1.2	Adesão à Terapêutica	11
1.2.1	Métodos para avaliar a adesão à terapêutica	14
1.2.2	Os fatores associados	17
2	Objetivos	19
3	Material e Métodos	21
3.1	Descrição Geral do Estudo	21
3.1.1	Escala de Satisfação com o Suporte Social (ESSS)	23
3.1.2	Escala de Ansiedade, Depressão e Stress (EADS)	24
3.1.3	Questionário da Perceção da Doença	26
3.1.4	Questionário de Crenças Acerca da Medicação	28
3.1.5	Variáveis em estudo	30
3.2	Análise Estatística	30
4	Modelo Linear Generalizado	33
4.1	Introdução	33
4.1.1	Valor Médio e Variância	35
4.1.2	Extensão do ML ao MLG	36
4.2	Regressão Logística	37

4.2.1	O Modelo	37
4.2.2	Coeficientes do Modelo	43
4.2.3	Estimação dos Parâmetros	46
4.2.4	Distribuições Assintóticas	48
4.2.5	Significância e Qualidade do Modelo de Regressão Logística	50
4.2.6	Pseudo- R^2	57
4.2.7	Capacidade Preditiva do Modelo	59
4.2.8	Intervalos de Confiança	64
5	Análise de Correspondências	65
5.1	Introdução	65
5.2	Análise de Correspondências	67
5.2.1	Matriz Inicial de Dados ou Matriz de Partida	68
5.2.2	Matriz de Correspondências	69
5.2.3	Massas ou Pesos	71
5.2.4	Perfis	73
5.2.5	Nuvens de Perfis	76
5.2.6	Centro de gravidade ou centróide	77
5.2.7	Distância	78
5.2.8	Inércia	80
5.2.9	Matriz de Covariâncias	83
5.2.10	Projeção de Indivíduos e Categorias. Princípio Baricêntrico	84
5.2.11	Decomposição em Valores Singulares	88
5.2.12	Aproximação da Matriz de Correspondências	90
5.2.13	Projeção de Indivíduos e Variáveis Suplementares	91
5.2.14	Número de Eixos a Reter	93
5.2.15	Contribuições	94
5.2.16	Algoritmo de Análise de Correspondências	96

5.3	Análise de Correspondências Múltiplas	98
5.3.1	Introdução	98
5.3.2	Matriz Concatenada	99
5.3.3	Matriz Disjuntiva Completa	104
5.3.4	Matriz de Burt	107
5.3.5	Inércias Ajustadas	110
5.3.6	Análise Classificatória como Auxiliar na Interpretação da ACM . .	112
6	Resultados	117
6.1	Análise Exploratória	117
6.1.1	Fatores Relacionados com o Doente	117
6.1.2	Fatores Relacionados com a Situação Clínica	132
6.1.3	Fatores Relacionados com as Crenças sobre a Doença e os Fármacos	135
6.2	Análise Bivariada	140
6.2.1	Fatores Relacionados com o Doente	141
6.2.2	Fatores Relacionados com a Situação Clínica	143
6.2.3	Fatores Relacionados com as Crenças sobre a Doença e os Fármacos	144
6.3	Análise de Regressão Logística	144
6.3.1	Análise Preliminar	144
6.3.2	Modelo de Regressão Logística Múltipla	149
6.4	Análise de Correspondências Múltiplas	155
6.4.1	Categorização das Variáveis de Interesse	155
6.4.2	Aplicação da ACM	157
7	Discussão e Problemas em Aberto	167
A	Referências Bibliográficas	173
B	Anexos	189

Lista de Figuras

1.1	Um olhar sobre o número de pessoas que vivem com HIV no mundo, 2011.	5
1.2	Casos de infecção VIH (1983-2012): distribuição por ano de diagnóstico e notificação.	6
3.1	Variáveis em estudo.	30
4.1	Estrutura do modelo de Regressão Logística.	38
4.2	Representação gráfica da função <i>logit</i>	40
4.3	<i>Confusion Matrix</i> para um modelo de regressão logística	60
4.4	Curva ROC.	63
5.1	Três conceitos geométricos básicos em Análise de Correspondências.	71
5.2	Quatro conceitos geométricos derivados em Análise de Correspondências.	72
5.3	Plano que contém os indivíduos caracterizados pelas categorias 1, 2 e 3.	74
5.4	Matriz ampliada da matriz K	92
5.5	Projeção ortogonal do elemento i sobre o eixo α , $pr_{\alpha(i)}$	95
5.6	Esquema geral da AC.	97
5.7	Exemplo de uma tabela tridimensional.	99
5.8	Análise de Correspondências Simples entre A e B.	100
5.9	Análise de Correspondências Simples entre A e C.	101
5.10	Análise de Correspondências Simples entre A e D.	102
5.11	Matriz disjuntiva completa genérica.	105
5.12	Conversão de uma tabela para a forma disjuntiva completa.	106

5.13	Análise de Correspondências Múltiplas via matriz disjuntiva completa.	106
5.14	Matriz de Burt genérica.	108
5.15	Conversão de uma tabela disjuntiva completa na matriz de Burt.	108
5.16	Análise de Correspondências Múltiplas via matriz de Burt.	109
5.17	Análise de Correspondências Múltiplas via matriz disjuntiva completa A vs B.	111
5.18	Réplica do gráfico apresentado na figura 5.13.	113
5.19	Dendrograma obtido através do Método da Ligação Completa aplicado à matriz de distâncias dada pelas coordenadas das categorias, no mapa apresentado na figura 5.18	114
5.20	Mapa ACM com grupos (de categorias) identificados.	115
6.1	Densidade de Kernel e histograma (à esquerda) e Box-plot (à direita) das idades dos participantes.	119
6.2	QQ-plot normal da distribuição das idades.	119
6.3	Gráfico de barras das categorias da variável idade.	120
6.4	Densidade de Kernel das idades por género (à esquerda), boxplots das idades dos indivíduos por género (ao meio) e Boxplots das idades dos in- divíduos por género com intervalos de confiança a 95% para a mediana (à direita).	121
6.5	QQ-plot normal da distribuição das idades para o género masculino (à esquerda) e para o género feminino (à direita).	122
6.6	Box-plots das pontuações (à esquerda) e box-plot das pontuações com in- tervalos de confiança (à direita), das três escalas da EADS.	127
6.7	Relação entre as variáveis Alcoolismo e Binge Drinking	131
6.8	Gráfico de barras relativos à classificação CDC nos três momentos avaliados.	134
6.9	Box-plots das pontuações das subescalas do <i>IPQ-R</i>	138
6.10	Box-plots das subescalas do instrumento BMQ.	140
6.11	Comparação entre os grupos dos participantes aderentes e não aderentes à terapêutica no que se refere ao facto de morarem ou não sozinhos	141
6.12	Comparação entre os grupos dos participantes aderentes e não aderentes à terapêutica no que se refere ao facto de serem ou não <i>binge drinkers</i>	142

6.13	Comparação da pontuação (score) média entre aderentes e não aderentes para a dimensão ansiedade, avaliada pela EADS	142
6.14	Comparação entre a percentagem doentes aderentes e não aderentes à terapêutica no que se refere à recomendação sobre TAR feita pelo médico . .	143
6.15	Comparação entre a percentagem de doentes aderentes e não aderentes à terapêutica no que se refere à recomendação de adesão à TAR feita pelo médico	144
6.16	Comparação das pontuações (scores) médias entre aderentes e não aderentes para as dimensões identidade, controlo de tratamento e representação emocional, avaliadas pelo (<i>IPQ-R</i>).	145
6.17	Comparação da pontuação (score) média entre aderentes e não aderentes para a dimensão preocupações, avaliada pelo BMQ.	145
6.18	Comparação da pontuação (score) média entre aderentes e não aderentes para a dimensão preocupações, avaliada pelo BMQ.	146
6.19	Gráficos dos resíduos da deviance, dos resíduos do qui-quadrado de Pearson e dos resíduos padronizados	151
6.20	Gráficos dos resíduos da deviance <i>vs</i> valores ajustados, dos resíduos do qui-quadrado de Pearson <i>vs</i> valores ajustados e dos resíduos padronizados <i>vs</i> valores ajustados	152
6.21	QQ-plot Normal dos resíduos do modelo	153
6.22	Curva ROC do modelo	154
6.23	Acertos, sensibilidade, especificidade e desempenho do modelo	154
6.24	Gráfico de barras das categorias das subescalas (variáveis) ansiedade e depressão.	156
6.25	Gráfico de barras das categorias das subescalas (variáveis) identidade, controlo de tratamento e representação emocional, respetivamente.	157
6.26	Gráfico de barras das categorias da subescala (variável) preocupações. . . .	157
6.27	Representação gráfica das duas primeiras dimensões da análise de correspondências múltiplas considerando as variáveis incluídas no modelo inicial de regressão logística múltipla	159
6.28	Representação gráfica das duas primeiras dimensões da análise de correspondências múltiplas com <i>zoom</i> em redor da categoria referente aos indivíduos aderentes (AACTG0)	160

6.29	Representação gráfica das duas primeiras dimensões da análise de correspondências múltiplas com <i>zoom</i> em redor da categoria referente aos indivíduos não aderentes (AACTG1)	161
6.30	Dendrograma das coordenadas das duas primeiras dimensões da análise de correspondências múltiplas cujo gráfico se apresenta na figura 6.27	162
6.31	Replicação do gráfico resultante da ACM.	162
6.32	Representação gráfica das duas primeiras dimensões da análise de correspondências múltiplas considerando as variáveis incluídas no modelo final de regressão logística múltipla	164
6.33	Dendrograma de apoio ao resultado da ACM ilustrado no gráfico 6.32 . . .	164

Lista de Tabelas

1.1	Horizonte temporal dos avanços e aprovações no tratamento do VIH na Europa (adptado de Fernandes (2011)).	7
1.2	Descrição dos métodos diretos para medição da adesão à terapêutica. . . .	15
1.3	Descrição dos métodos indiretos para medição da adesão à terapêutica. . .	16
3.1	Classificação dos níveis de adesão à terapêutica anti-retrovírica, em função das respostas dadas para os diferentes momentos ou períodos de tempo avaliados	22
3.2	Alfa de Cronbach para as subescalas do ESSS	24
3.3	Alfa de Cronbach para as subescalas do EADS	26
3.4	Alfa de Cronbach para as subescalas do <i>IPQ-R</i>	28
4.1	Capacidade preditiva do modelo de acordo com a sensibilidade e especificidade	61
4.2	Valores indicativos da área sob a curva ROC	63
5.1	Resumo sobre a nuvem de pontos	81
5.2	Matriz concatenada	100
6.1	Caracterização demográfica da amostra	118
6.2	Caracterização dos co-habitantes dos participantes	123
6.3	Caracterização escolar e profissional dos participantes	124
6.4	Alfa de Cronbach para a escala ESSS no presente estudo	124
6.5	Alfa de Cronbach para as subescalas subescala da ESSS no presente estudo	124
6.6	Classificação da consistência interna de acordo com o alfa de Cronbach . .	125

6.7	Correlação de Spearman entre as subescalas da ESSS	125
6.8	Alfa de Cronbach para a escala EADS no presente estudo	126
6.9	Alfa de Cronbach para as subescalas do EADS no presente estudo	126
6.10	Correlação de Spearman entre as subescalas da EADS	126
6.11	Caracterização dos hábitos de exercício físico dos participantes	128
6.12	Caracterização dos hábitos de exercício físico dos participantes	129
6.13	Classificação dos hábitos alcoólicos de acordo com o número médio de be- bidas ingeridas por dia, nos últimos 30 dias	130
6.14	Classificação de <i>binge drinker</i>	130
6.15	Caracterização do consumo de cigarros e drogas	131
6.16	Tratamento sob metadona	132
6.17	Modo de transmissão do vírus	132
6.18	Tempo até iniciar TAR, sob TAR e de seguimento (em meses)	133
6.19	Classificação CDC dos indivíduos nos três momentos avaliados	134
6.20	Alfa de Cronbach para as subescalas do IPQ-R no presente estudo	135
6.21	Correlação de Spearman entre as subescalas do <i>IPQ-R</i>	136
6.22	Alfa de Cronbach para a escala BMQ no presente estudo	138
6.23	Alfa de Cronbach para as subescalas do BMQ no presente estudo	139
6.24	Correlação de Spearman entre as subescalas do BMQ	139
6.25	Variáveis que se verificaram relacionadas com a não-adesão, aquando apli- cação do teste do qui-quadrado (ou teste exato de Fisher) ou teste de Mann-Whitney	141
6.26	Análise univariada dos fatores da não-adesão	149
6.27	Análise de regressão logística multivariada, com <i>stepwise</i> , dos fatores asso- ciados à não-adesão	150
B.1	Grelha de correção da Escala de Satisfação com o Suporte Social (ESSS) .	189
B.2	Grelha de correção da Escala de Ansiedade, Depressão e Stress (EADS) . .	190
B.3	Grelha de correção da subescala identidade da escala IPQ-R	191

B.4	Grelha de correção das restantes subescalas da escala IPQ-R	192
B.5	Grelha de correção da escala de crenças acerca dos fármacos	193

Lista de acrónimos e abreviaturas

AACTG - *AIDS Adult Clinical Trial Group*
AC - Análise de Correspondências
ACP- Análise em Componentes Principais
ACTG - *AIDS Clinical Trials Group*
ACM - Análise de Correspondências Múltiplas
ADN - Ácido desoxirribonucleico
ARN - Ácido Ribonucleico
BMQ - *Beliefs about Medicine Questionnaire*
CV - Carga vírica
DASS - *Depression Anxiety Stress Scale*
DVS - Decomposição em Valores Singulares
EADS - Escala de Ansiedade, Depressão e Stress
EDEM - *Electronic Drug Exposure Monitor*
ESSS - Escala de satisfação com o Suporte Social
FDA - *Food and Drug Administration*
FFS - Formas Farmacêuticas Sólidas
f.d.p. - Função Densidade de Probabilidade
f.m.p. - Função Massa de Probabilidade
GLM - *Generalized Linear Models*
GRID - *Gay-related Immune Deficiency*
HAART - Terapêutica Anti-retrovírica de Alta Eficácia (*Highly Active Antiretroviral Therapy*)
HSM - Hospital de Santa Maria
IC - Intervalo de Confiança
ICN - *International Council of Nurses*
IMP - Instituto de Medicina Preventiva
IPQ - *Illness Perception Questionnaire*
IPQ-R - *Revised Illness Perception Questionnaire*
MC - Matriz Concatenada MDC - Matriz Disjuntiva Completa MEMS - *Medication Event*

Monitoring System

ML - Modelo Linear

MLM - Modelo Linear Misto

NIH - National Institute of Health

OMS - Organização Mundial de Saúde

ROC - *Receiver Operating Characteristic*

SIDA - Síndrome da Imunodeficiência Adquirida

TAR - Terapêutica anti-retrovírica

TARc - Terapêutica anti-retrovírica combinada

TB - Tuberculose

VIH - Vírus da Imunodeficiência Humana

VIH-1 - Vírus da Imunodeficiência Humana tipo 1

VIH-2 - Vírus da Imunodeficiência Humana tipo 2

Estrutura da tese

Com vista a atingir os objetivos propostos, no capítulo 1, apresentar-se-á uma breve introdução sobre a infeção pelo vírus da imunodeficiência humana. Em particular, vai ser dado especial interesse ao tratamento, bem como a adesão ao mesmo, nomeadamente, aos vários métodos disponíveis de avaliar a adesão. Por fim, serão referidos os fatores que são comumente apontados como preditores de não adesão.

No capítulo dois serão apresentados os objetivos gerais do presente trabalho.

No terceiro capítulo ir-se-á descrever, resumidamente, o estudo conduzido pelo Instituto de Medicina Preventiva, o qual permitiu a realização deste trabalho. Vão ser descritos os instrumentos de recolha e a análise estatística efetuada.

Os quarto e quintos capítulos ir-se-ão focar na apresentação de conceitos teóricos sobre os modelos estatísticos utilizados, nomeadamente, o modelo de Regressão Linear Simples e o modelo de Análise de Correspondências Múltiplas, respetivamente.

No capítulo seis será efetuada a análise ao conjunto de dados em estudo. Começar-se-á por realizar uma análise exploratória e, posteriormente, serão aplicadas as metodologias referidas nos dois capítulos anteriores, com vista a ir ao encontro dos objetivos inicialmente propostos e acima apresentados.

Finalmente, no capítulo último, sete, será efetuada a discussão geral dos resultados expostos no capítulo anterior. Vão ser igualmente mencionadas algumas considerações finais e possíveis problemas em aberto.

Capítulo 1

Introdução

1.1 A infecção VIH/SIDA

A SIDA é o Síndrome da Imunodeficiência Adquirida, sendo provocada pelo vírus da Imunodeficiência Humana (VIH), cujo modo de transmissão é através de relações sexuais, injeção ou transfusão de sangue ou de produtos sanguíneos contaminados, partilha de material de injeção não esterilizado, que foi previamente usado por alguém que está infetado, e transmissão materno-fetal (Infectious Disease Prevention and Control, 2012).

O atual tratamento para a infecção pelo VIH tem contribuído para uma redução drástica da morbilidade e mortalidade em indivíduos infetados pelo VIH, principalmente, a partir de 1996, com o uso da terapêutica anti-retroviral combinada (TARc), designada também por terapêutica tripla (*Highly Active Antiretroviral Therapy (HAART)*). Desta forma, a TARc transformou a infecção pelo VIH de uma doença fatal para uma doença crónica, com maior tempo de sobrevivência dos indivíduos seropositivos.

Os resultados com a TARc mostram a eficácia terapêutica e, consequentemente, revelam a importância que os prestadores de cuidados de saúde têm para a obtenção contínua de melhores resultados. Estudos mostram que a não-adesão, ocorre em algum grau, tanto em países desenvolvidos como em países subdesenvolvidos, sendo que a taxa média é de 50%.

A não-adesão aos medicamentos anti-retrovirais é, hoje em dia, considerada como um dos principais determinantes da efetividade do tratamento, no plano individual. Sem um nível de adesão adequado, os fármacos anti-retrovirais não são mantidos em concentrações suficientes para suprimir a replicação do vírus e impedir a diminuição da carga viral. Além de estar associada a uma má resposta virológica a curto prazo, a fraca adesão ao regime anti-retrovírico, acelera a resistência ao fármaco, uma vez que, quando se omite doses,

podem desenvolver-se estirpes de VIH resistentes aos medicamentos prescritos. Portanto, a identificação dos fatores que reduzem a adesão à terapêutica assume um papel fulcral para a supressão prolongada da carga viral.

Os fatores críticos que influenciam a adesão recaem, essencialmente, sobre 4 grupos: (1) Fatores relacionados com o doente, tais como o uso de drogas, consumo de álcool, idade, sexo ou etnia; (2) Regime terapêutico, tais como a complexidade do plano medicamentoso, o número de comprimidos ou necessidades alimentares; (3) A relação doente-prestador de cuidados de saúde (enfermeiros e/ou médicos); (4) Sistema de atendimento.

Deste modo, tendo em conta que a infeção pelo vírus da imunodeficiência humana se tornou uma pandemia a nível global, torna-se fundamental determinar quais os fatores que influenciam a adesão à terapêutica anti-retrovírica, permitindo incrementar estratégias de intervenção junto dos serviços de saúde e avaliá-las a partir da compreensão dos fatores associados à adesão.

"Fui diagnosticado para o VIH em 1998. Durante cerca de 7 anos, como as minhas defesas estavam boas e a carga viral não muito alta, não precisei de fazer a medicação. Iniciei a terapêutica em março de 2006. psicologicamente, foi um grande desafio que enfrentei, mas que superei com o tempo e, hoje, abril de 2013, tomo apenas um comprimido diário. Neste momento, e graças à medicação, as minhas defesas estão ótimas, a carga viral indetetável e levo uma vida perfeitamente normal e ativa"

Retirado de Santos (2013)

Até 31 de dezembro de 2011, foram notificados perto de 2.5 milhões de novos casos de infeção VIH/SIDA em todo o mundo, juntando-se assim aos já cerca de 31.5 milhões infetados até então. Assim, até ao final de 2011, encontravam-se diagnosticados, aproximadamente 34 milhões de infetados pelo VIH, dos quais, 8 milhões encontram-se sob TAR (figura 1.1).

Em Portugal, os últimos dados remontam ao ano de 2012, durante o qual foram notificados 1707 casos de infeção pelo vírus da imunodeficiência humana, referente a 1551 novos casos e a 156 casos em que se verificou evolução do estágio. Entre estes casos, 776 (45,5%) foram diagnosticados entre 1 de janeiro e 31 de dezembro de 2012 e os restantes 931 (54,5%) foram detetados em anos anteriores, porém, a sua sinalização apenas foi efetuada em 2012 (Infeção VIH/SIDA: a situação em Portugal a 31 de Dezembro de 2012).

No final de 2012 o total acumulado de casos de infeção VIH, declarados em Portugal desde 1985, ascendia a 42 580, dos quais 17 373 se encontravam no estágio SIDA. A distribuição dos casos notificados por ano de diagnóstico e por ano de notificação é apresentada na figura 1.2.

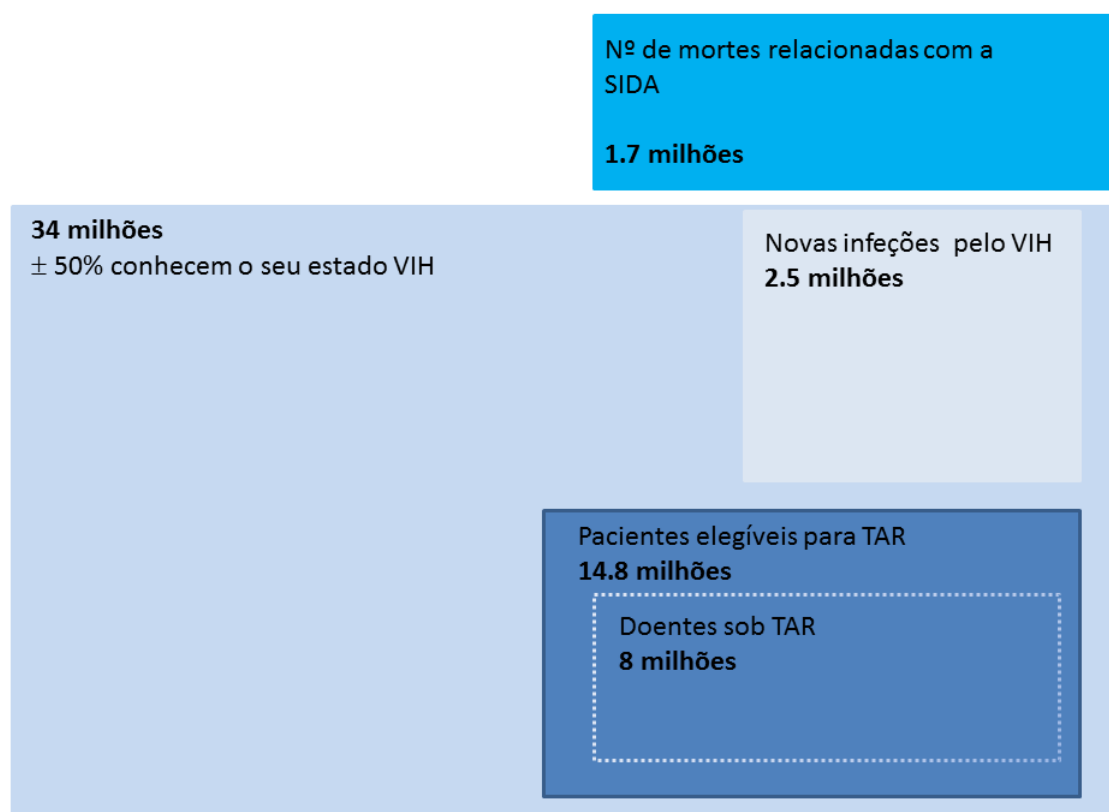


Figura 1.1: Um olhar sobre o número de pessoas que vivem com HIV no mundo, 2011 (adaptado de *World AIDS Day Report 2012*).

Uma análise rápida revela que os casos mais antigos de VIH registados remontam ao ano de 1983 e que 2000 é o ano em que se contabiliza o maior número de casos diagnosticados (2795), bem como o maior número de casos notificados (3674).

De referir que, em 2005, passou a ser obrigatória a notificação dos casos de infeção pelo VIH, o que conduziu a que nos anos subsequentes fossem notificados casos diagnosticados em anos anteriores, facto que ainda se verifica hoje em dia.

Dos casos diagnosticados em 2012, verificou-se que 319 (41,1%) referem residência no distrito de Lisboa e 549 (70,7%) eram do sexo masculino verificando-se um ratio homem/mulher de, aproximadamente, 2,4. A maioria dos casos, 548 (70,6%), foi registada em indivíduos de nacionalidade portuguesa.

Em 2012 foram verificados 496 novos casos no estágio SIDA, dos quais 247 foram detetados no próprio ano, maioritariamente, no distrito de Lisboa (39,3%). 184 (74,1%) eram do sexo masculino, traduzindo-se num ratio homem/mulher de, aproximadamente, 2,9.

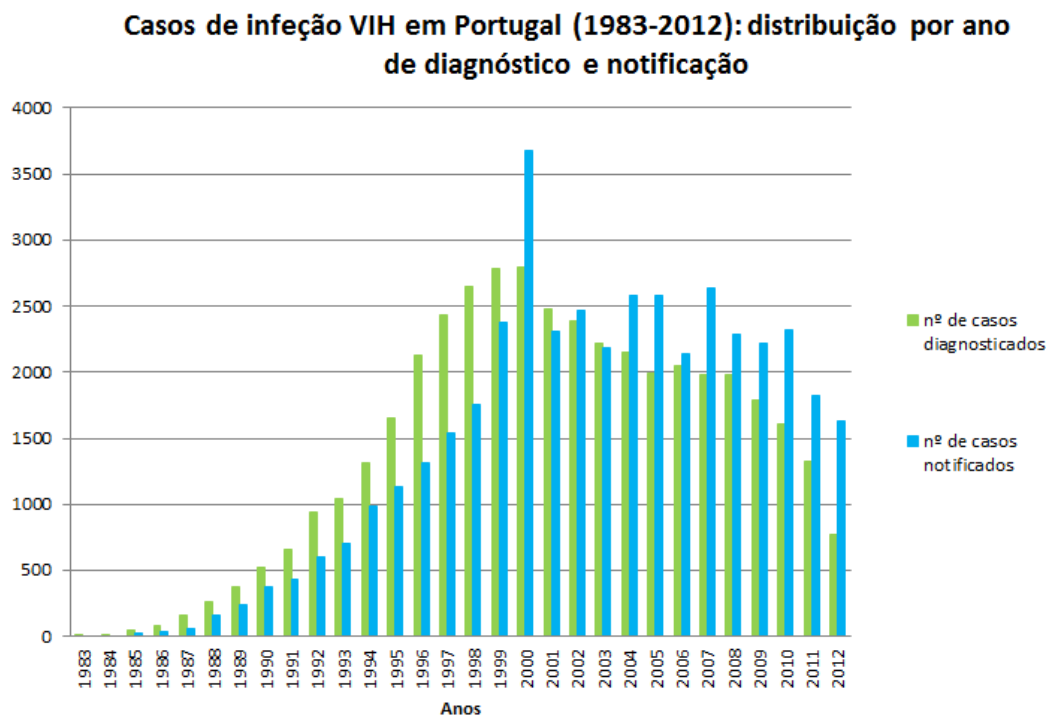


Figura 1.2: Casos de infeção VIH (1983-2012): distribuição por ano de diagnóstico e notificação (adaptado de Infeção VIH/SIDA: a situação em Portugal a 31 de Dezembro de 2012).

1.1.1 Terapêutica Anti-Retrovítica (TAR)

A terapêutica anti-retrovítica, também denominada de terapêutica anti-retroviral, consiste no tratamento para a eliminação de retrovírus do organismo. Porém, frequentemente, este termo é utilizado para descrever o tratamento contra o VIH, o retrovírus mais famoso. Tem como objetivo reduzir a carga viral e repor a contagem de células CD4 e a função imunitária. Passados 32 anos, a infeção pelo VIH passou de uma doença fatal para uma doença aguda crónica. Durante este longo período, o desenvolvimento da terapêutica anti-retrovítica (TAR) tornou-se num dos avanços mais importantes na história da medicina, passando de uma monoterapia com a Zidovudina para o número de opções anti-retrovirais disponíveis hoje em dia (tabela 1.1).

A introdução dos inibidores não-nucleósidos da transcriptase reversa e dos inibidores da protease, nos regimes terapêuticos anti-retrovirais em 1995, correspondeu ao início da era da HAART, a qual consiste na combinação de vários medicamentos anti-retrovirais com a finalidade de retardar a progressão do vírus e tem como principais objetivos (Guerreiro *et al.*, 2012):

- Reduzir a mortalidade e morbilidade associadas à infeção pelo VIH;

Tabela 1.1: Horizonte temporal dos avanços e aprovações no tratamento do VIH na Europa (adaptado de Fernandes (2011) e Pomerantz and Horn (2013))

Ano	Avanços	Nome do fármaco (ano de aprovação)	Classe anti-retrovírica
1981	* Reconhecimento de uma nova doença denominada SIDA	—	
1982	* Reconhecimento da doença prejudicada pela falta de testes de diagnóstico	—	
1983	* Isolamento do VIH-1	—	
1985	* Isolamento do VIH-2	—	
1987	* Aprovação do 1º fármaco contra VIH-1 pelo FDA ¹ , disponível nos E.U.A	1990, Zidovudina	INTR
1991	* Novo INTR pronto para aprovação pelo FDA	1992, Didanosina	INTR
1994	* Detetadas limitações da monoterapia * Estudos clínicos revelam que o uso da Zidovudina em grávidas reduz a transmissão do VIH-1 ao feto	Zalcitabina	INTR
1995	* Compreensão inicial da dinâmica viral do VIH-1 * Lançamento dos primeiros inibidores da protease (IP) na <i>HAART</i> - 1996 * Divulgação do teste de carga viral plasmática	1996, Estavudina 1996, Lamivudina	INTR INTR
—	Primeiro INNTR aprovado pela FDA * Redução contínua da incidência da SIDA e morte nos países desenvolvidos, principalmente, devido à TARc	1996, Saquinavir - HCG	IP
1997	* Reconhecimento das limitações da TARc: desenvolvimento da resistência e efeitos adversos da terapêutica	1996, Ritonavir 1996, Indinavir	IP IP
1998	* Reconhecimento da latência do VIH * Restauração da função imunitária com a TARc * Descoberta de que o tratamento inicial com efavirenz é, pelo menos, tão eficaz como os IP tão eficaz como os IP	1998, Nevirapina 1998, Nelfinavir 1997, Delavirdina 1998, Lamivudina + Zidovudina 1998, Saquinavir	INNTR IP INNTR 2NNTR IP
1999	* Em doentes com falência virológica, verificou-se que a adição de 2 fármacos resulta numa supressão maior do que a adição de apenas 1 fármaco * Utilização de novas combinações de fármacos adicionados à TARc, numa tentativa de erradicar o VIH-1 nos indivíduos seropositivos	1999, Efavirenz 1999, Abacavir	INNTR INTR
2000	* Maior compreensão dos reservatórios, tornando a erradicação do VIH inatingível com os fármacos atuais * Ritonavir impulsionou a classe dos fármacos IP * Testes de resistência genotípica podem orientar a seleção do melhor regime	2000, Amprenavir 2000, Didanosina (EC) 2000, Abacavir + Lamivudina + Zidovudina	IP INTR 3INTR
2001	* Inibidores da transcriptase reversa análogos (especialmente a Estavudina) causam muitos dos efeitos adversos a longo prazo da TARc (neuropatia, lipodistrofia, acidose láctica). Estes efeitos não se verificam com o Tenofovir e na terapêutica com Abacavir	2001, Lopinavir + Ritonavir	PI

Ano	Avanços	Nome do fármaco (ano de aprovação)	Classe anti-retrovírica
2002	* Prevenção da transmissão vertical ² nos países desenvolvidos * Novas diretrizes para início da TARc (CD4 < 350 células/ml)	2002, Tenofovir	INTR
2003	* Introdução dos inibidores de fusão * Necessidade de traduzir o progresso para os países desenvolvidos - Declaração do Millennium	2003, Enfuvirtida 2003, Emtricitabina 2003, Fosamprenavir	IE INTR IP
—	* Disponibilização de esquemas terapêuticos com reduzida carga de comprimidos e frequência de administração;	2004, Atazanavir 2004, Abacavir + Lamivudina	IP 2INTR
2006	vários tratamentos de combinações de dose fixa * Primeira tentativa de criação de vacina contra o VIH	2005, Tenofovir + Emtricitabina 2005, Tripanavir	2INTR IP
2006	* Vários fármacos anti-retrovíricos nas classes existentes e numa nova classe com nova atividade contra a resistência aos fármacos são aprovados;	2007, Darunavir	IP
—	uso de, pelo menos, 2 e, por vezes, 3 fármacos ativos na prática clínica, produzindo elevadas taxas de supressão viral, inclusive em doentes que não se encontram no início do tratamento	2007, Efavirenz + Emtricitabina + Tenofovir	INNTR + 2INTR
2008	* Aprovação da combinação de dose fixa, um comprimido, uma vez por dia	2007, Maraviroc 2007, Raltegravir	IE II
2008	* Maior conhecimento dos benefícios do início precoce e risco da interrupção do tratamento com contagens elevadas de células CD4		
—	* Diretrizes recomendam TAR ao longo da vida dos doentes assintomáticos e início de TARc precoce (CD4 < 500 células/ml)		
2010	* Crescente interesse em vacinas "terapêuticas", com o objetivo de melhorar a resposta imunitária, para atuar sinergicamente com a terapêutica anti-retrovírica		
2011	* Declaração Política das Nações Unidas sobre HIV/SIDA: Intensificação dos esforços para eliminar o HIV/SIDA		

¹FDA - *Food and Drug Administration*.

²Transmissão de uma infeção ou doença da mãe para o feto, isto é, transmissão materno-infantil.

¹INTR - Inibidores não nucleósidos/nucleótidos de transcriptase reversa; INNTR - Inibidores não-nucleósidos da transcriptase reversa; IP - Inibidores da protease; IE - Inibidores de entrada; II - Inibidores de integração.

- Proporcionar uma qualidade de vida digna para os doentes;
- Restaurar e/ou preservar, na medida do possível, a função imunitária;
- Maximizar a supressão da replicação do vírus, isto é, reduzir o máximo e durante o maior tempo possível, a carga viral, de forma a retardar a progressão da doença;
- Prevenir/reduzir a transmissão viral.

O objetivo desta terapêutica visa prolongar e melhorar a qualidade de vida do doente, a qual é consequência de uma supressão máxima e duradoura da carga vírica. Deve,

portanto, tratar-se de uma estratégia precoce, de modo a manter a imunidade e/ou a retardar a progressiva destruição do sistema imunológico. Esta destruição é responsável pela infecções oportunistas que advêm da infecção com o vírus VIH. De notar que estas infecções são as principais responsáveis pela morbilidade e mortalidade associada ao vírus. A utilização de regimes terapêuticos fortes está diretamente associada à inibição da multiplicação vírica e, em geral, à recuperação da resposta imunológica, com a subida entre 100 a 200 linfócitos T $CD4/mm^3$ por ano. O insucesso terapêutico é caracterizado por 3 fatores: persistência da carga vírica; redução de células $CD4$ em 30%; ocorrência de complicações clínicas.

São vários os motivos que podem conduzir à falência de um regime terapêutico², tais como:

- **Fatores associados com o doente:** Idade, data de início da terapêutica, carga vírica (CV) e número de linfócitos T CD antes da submissão à terapêutica, doenças oportunistas prévias, co-morbilidades, uso de drogas, resistências a fármacos anteriores, uso anterior de TAR, consumo de álcool;
- **Potência do regime terapêutico:** Uso de uma combinação com menos de três fármacos ativos no indivíduo;
- **Má assimilação dos medicamentos pelo organismo:** Embora os medicamentos possam estar a ser tomados como administrado, a sua assimilação pelo organismo pode não correr como expectável. Por exemplo, para alguns fármacos pode ser necessário um ajuste na dose em função do peso corporal. Noutras situações, a absorção do fármaco pode ser alterada aquando a ingestão de medicamentos;
- **Resistência prévia a medicamentos administrados anteriormente:** Ao juntar-se novos fármacos aos que já estavam a ser administrados, aumenta-se o risco de ocorrência de resistências. Quando se está infetado com uma estirpe de vírus já resistente, por exemplo, ao AZT, o uso deste fármaco torna-se totalmente ineficaz;
- **Baixo nível de adesão ou adesão não regular ao tratamento:** A adesão é fundamental para o sucesso da terapêutica e a adesão ótima é tão eficaz como a introdução de um novo medicamento. Deste modo, a resistência e a adesão estão intimamente relacionadas. Quando um doente se esquece ou se atrasa na toma

²A definição de falência terapêutica baseia-se nos critérios clínicos, imunológicos e virológicos. Diversos elementos podem ser usados para definir falência. O mais comum, na infecção VIH, é a progressão clínica da doença, isto é, o surgir de uma nova infecção oportunista. Aparição de uma nova infecção oportunista ou de uma patologia maligna ligada à progressão da infecção VIH

da medicação, aumenta o risco do desenvolvimento de resistências. Isto acontece porque, os níveis dos medicamentos descem abaixo do nível mínimo de segurança necessário para controlar o vírus (Santos, 2008).

De notar que a resistência a medicamentos apenas se desenvolve quando se está a fazer um tratamento que não suprime completamente a infeção. Pode desenvolver-se resistência a tratamentos para o vírus VIH ou da hepatite B, a infeções bacterianas como a Tuberculose, entre outras. Tal ocorre quando os tratamentos não são suficientemente potentes para impedir a replicação do vírus. É por esta razão que, no caso do VIH, normalmente, é necessária a combinação de, pelo menos, três medicamentos.

As resistências são originadas mais rapidamente quando a infeção apresenta uma taxa de replicação elevada. No caso dos doentes seropositivos, todos os dias são produzidos biliões de novas cópias do vírus. Para os indivíduos que se encontram sujeitos à terapêutica anti-retrovírica, quanto mais elevada for a carga viral, maior é a probabilidade de desenvolverem resistências (Santos, 2008). Por esta razão, torna-se crucial que, aquando o início da terapêutica, a CV se encontre o mais baixa possível.

Atualmentes existem 5 tipos de fármacos anti-retrovíricos, que atrasam o desenvolvimento da doença através do VIH:

- **Inibidores nucleósidos/nucleótidos da transcriptase reversa (INTR)** - O seu alvo de ação é a enzima transcriptase reversa, atuando como substratos alternativos, competindo com os nucleósidos fisiológicos, diferindo destes numa simples alteração numa molécula de açúcar. Esta clonagem impede a replicação vírica, evitando a infeção de novos linfócitos.
- **Inibidores não-nucleósidos da transcriptase reversa (INNTR)** - Ligam-se à transcriptase reversa e, após estarem unidos à enzima, afetam a atividade da mesma, restringindo a sua mobilidade num ponto crítico, impedindo a sua função. Deste modo, a enzima fica impossibilitada de interagir com o ARN vírico, inibindo a produção de ADN vírico.
- **Inibidores da protease (IP)** - Atuam sobre a enzima protease, impedindo a maturação dos organismos virais.
- **Inibidores de entrada (IE)** - Inibem a ligação do vírus aos recetores das células, evitando, desta forma, a sua entrada nas mesmas.
- **Inibidores de integrase(II)** - Impedem a integração do DNA viral no DNA humano.

1.2 Adesão à Terapêutica

A adesão ao regime terapêutico, adesão terapêutica ou ainda adesão ao tratamento são diferentes formas de referir a mesma entidade (Vermeire *et al.*, 2011).

Para a maioria das situações clínicas, o diagnóstico correto e tratamento médico eficaz são essenciais para a sobrevivência do doente bem como da qualidade de vida do mesmo. O obstáculo para a eficácia do tratamento é, no entanto, a incapacidade do doente em seguir as recomendações feitas pelo médico e profissionais de saúde. Porém, a não-adesão ao tratamento pode tomar muitas formas. Os conselhos dados pelos prestadores de cuidados de saúde são, frequentemente, mal entendidos, esquecidos ou ignorados. A não-adesão acarreta custos económicos bastantes elevados e, segundo DiMatteo (2004), os gastos anuais como consequência da não-adesão ascendem ao valor estimado de centenas de biliões de dólares. Os custos auferidos com hospitalização relacionada com a mesma, foi estimada, por Sullivan (1990), em \$13.35 biliões/ano.

Deste modo, a adesão à terapêutica assume um papel de particular importância nos doentes portadores de doenças crónicas, sendo que, a não-adesão é, hoje em dia, considerada um problema de saúde pública.

Entende-se adesão como o grau de conformidade entre as recomendações dos profissionais de saúde e o comportamento do doente relativamente ao regime terapêutico proposto (Haynes *et al.*, 2008). Este conceito pode também ser visto como um conjunto de comportamentos, tais como: tomar a medicação, seguir dietas ou proceder a mudanças de hábitos de vida que vão de encontro ao plano terapêutico prescrito (Almeida *et al.*, 2007).

Os participantes no congresso da WHO, em Junho de 2001, definiram a adesão como "o grau segundo o qual o doente segue as instruções médicas" (Sabaté, 2001).

Sackett e Haynes, em 1976, introduziram a expressão "adesão ao tratamento" na literatura médica, como sendo "o ponto em que o comportamento de uma pessoa que toma medicamentos, segue uma dieta recomendada ou muda de estilo de vida, coincide com conselhos médicos ou de saúde". Segundo a visão destes dois autores, tomar a medicação é um fim e não um meio para tornar o tratamento eficaz. Este conceito acentava no cumprimento escrupuloso das indicações do médico e equipas de saúde, dando ênfase a uma relação paternalista entre estes e o doente, pelo que a componente de partilha de decisões era muito pouco valorizada. Esta noção de adesão não possuía quaisquer graus, isto é, o doente ou cumpria as indicações do médico ou não.

Um pouco mais tarde, foi desenvolvido o conceito de "aliança terapêutica" que descrevia a relação doente-médico como um processo de interação. A definição de "aliança

terapêutica", mais frequentemente, aceite é: "o ponto em que o indivíduo escolhe comportamentos que coincidem com uma prescrição clínica, o regime deve ser consensual, ou seja, alcançado através de negociações entre o profissional de saúde e o doente" (Lopes *et al.*, 2008). À semelhança da definição introduzida em 1976, esta definição de "aliança terapêutica" parece ignorar quer o resultado da partilha de decisões na saúde do doente, quer a existência de níveis de adesão.

Em 2003, a Organização Mundial de Saúde (OMS) definiu a adesão como sendo "a extensão em que o comportamento de um indivíduo - utilização de medicamentos e cuidados de saúde, alimentação e estilos de vida - está de acordo com as recomendações de um profissional de saúde (Sabaté, 2001). Com a introdução desta nova definição, passou a considerar-se níveis de adesão, tendo sido introduzido, em simultâneo, o conceito de concordância, isto é, a adesão como um meio para alcançar o sucesso terapêutico e não um fim.

De acordo com a Classificação Internacional para a Prática de Enfermagem (CIPE, do inglês ICN - *Internacional Council of nurses*) (2011), o termo adesão ao regime terapêutico é definido como uma "Ação auto-iniciada para promoção de bem-estar, recuperação e reabilitação, seguindo as orientações sem desvios, empenhado num conjunto de ações ou comportamentos. Cumpre o regime de tratamento, toma os medicamentos como prescrito, muda o comportamento para melhor, sinais de cura, procura os medicamentos na data indicada, interioriza o valor de um comportamento de saúde e obedece às instruções relativas ao tratamento. (Frequentemente associado ao apoio da família e de pessoas que são importantes para o cliente, conhecimento sobre os medicamentos e processo de doença, motivação do cliente, relação entre o profissional de saúde e o cliente)".

Recentemente, surge o termo "adesão parcial", que corresponde a situações em que o doente não manifesta uma adesão total. Quando um doente toma a dose, mas na hora incorreta e/ou a dose incorreta, diz-se que aderiu parcialmente à terapêutica. Esta adesão pode ser intencional ou não intencional. Enquanto a adesão intencional acontece numa fase mais aguda da doença (por exemplo, quando o doente se sente pior, toma a medicação e quando não têm qualquer sintoma da doença, ignora a medicação) a adesão não intencional é mais frequente nos idosos e crianças, sendo causada, maioritariamente, pelo esquecimento da dose prescrita, pela confusão nos esquemas terapêuticos, instruções não perceptíveis ou dificuldade na toma da medicação.

Na atualidade, uma percentagem significativa da população mundial enfrenta a necessidade de tomar medicação e de adotar medidas para controlar e tratar, tanto as patologias agudas como as crónicas.

A não-adesão pode ser vista como "universal" e está relacionada com diversos fatores

relativos ao tratamento, aos profissionais de saúde, à patologia e ao doente (Jordan, 2000, Paulo e Zanini, 1997).

Estima-se que, nos países desenvolvidos, o grau de adesão às terapêuticas associadas às doenças crónicas seja apenas de 50% (Dias *et al.*, 2011) e, nos países subdesenvolvidos, ou em vias de desenvolvimento, seja ainda menor, condicionando a economia, o bem-estar e a qualidade de vida da sociedade. A escassez de recursos e as desigualdades sociais no acesso aos cuidados de saúde transformaram a não-adesão num problema de grande dimensão, em determinados países (OMS, 2003). Alguns trabalhos realizados sobre esta temática, demonstram que 6 a 20% dos doentes não levantam a medicação nas farmácias e que 30 a 50% não cumprem o esquema proposto, atrasando ou omitindo doses (Sabaté, 2001). Entre as muitas razões indicadas pelos indivíduos para não cooperar no plano terapêutico, a mais comum é o esquecimento.

É, portanto, de capital interesse conhecer os fatores que influenciam a adesão, percebendo a associação entre variáveis internas e externas ao doente, bem como as variáveis relacionais. Desta forma, poder-se-ão conceber estratégias que promovam a adesão ao regima terapêutico, que a nível educacional, comportamental ou conjungando as duas dimensões.

Normalmente, o **doente aderente** é aquele que:

- Não falha qualquer toma da medicação;
- Não altera a dose prescrita pelo médico e/ou profissionais de saúde;
- Cumpre o esquema prescrito;
- Cumpre a relação dos medicamentos com os alimentos (por exemplo, tomar após as refeições).

Por contraposição, são manifestações de **não-adesão**:

- Não aviar os fármacos prescritos;
- Falhar doses;
- Não cumprir o esquema medicamentoso;
- Sobredosagem;
- Interromper a medicação antes do tempo;
- Tomar medicamentos com alimentos ou outra medicação quando contra indicados;

- Não comparecer às consultas;

No que concerne ao tratamento para o VIH/SIDA, encontram-se definidas 3 alicerces fundamentais para o sucesso terapêutico: Capacidade do doente no sistema de saúde; retenção/fidelização do doente ao sistema de assistência médica e adesão à terapêutica (Programa Nacional para a infeção VIH/SIDA, 2012).

1.2.1 Métodos para avaliar a adesão à terapêutica

Vários foram os estudos que nos últimos 30 anos se debruçaram sobre a temática da adesão na infeção VIH/SIDA. Estudos recentes mostraram que é necessário cumprir entre 90 a 95% do plano terapêutico, tendo em vista o sucesso terapêutico (referir o livro 10).

Não tem sido, no entanto, fácil estabelecer o método (mais) adequado para avaliar a adesão e o limiar ideal a assumir para esta patologia.

Tradicionalmente, os métodos aplicados são classificados como:

- Diretos:

- i) Determinação das concentrações plasmáticas de fármacos ou metabolitos;
- ii) Marcadores biológicos, como a contagem de linfócitos CD4 e a carga viral;

- **Indiretos:** Avaliam a adesão através das informações facultadas pelo doente ou pelo profissional de saúde. Permite, assim, inferir de forma indireta a adesão do doente à terapêutica:

- i) Auto-relato da adesão;
- ii) Avaliação clínica;
- iii) Contagem de forma farmacêuticas sólidas (FFS) ou, mais frequentemente, denominado de contagem de comprimidos;
- iv) Registo de dispensa dos medicamentos por parte dos serviços farmacêuticos;
- v) Sistemas de monitorização eletrónica.

Nas tabelas 1.2 e 1.3 descrevem-se resumidamente os métodos utilizados. É, contudo, importante referir que não existe um método assumido como referência, isto é, não existe um método que prevalece sobre os restantes. Apesar disso, no contexto de estudos epidemiológicos, o recurso ao auto-relato e à avaliação pelo médico são escolhas frequentes (Romeu *et al.* (2012), Remor *et al.* (2007)).

Tabela 1.2: Descrição dos métodos diretos para medição da adesão à terapêutica

Métodos diretos	
Método	Descrição
Determinação das concentrações plasmáticas de fármacos ou metabolitos	Baseia-se na determinação analítica do medicamento ou dos seus metabolitos no plasma. Ao contrário de outros métodos, consiste numa metodologia objetiva, permitindo inferir níveis de incumprimento ou não-adesão, superiores aos dos métodos indiretos. Tem, contudo, a desvantagem de ser mais dispendioso e sensível ao chamado "efeito de bata branca" ou "efeito escova de dentes", isto é, verifica-se um aumento do cumprimento da terapêutica nos dias imediatamente anteriores ao da consulta
Marcadores biológicos	Um marcador biológico é, por definição, uma substância não tóxica, estável e facilmente detetável e que pode ser adicionada ao fármaco. A deteção desta substância no sangue, ou na urina, reflete a presença do medicamento. De referir que, a utilização de marcadores biológicos é restrita à investigação e é, atualmente, alvo de grande controvérsia e de difícil aplicação

1.2.2.1 Instrumento *Adult AIDS Clinical Trials Group (AACTG)*

Com o objetivo de avaliar a adesão aos novos regimes terapêuticos, foi desenvolvido um instrumento, originalmente denominado de *Adult AIDS Clinical Trials Group (AACTG)* (Chesney, 2000), o qual procurou estratégias para avaliar a adesão e os principais motivos de não-adesão. Trata-se de um questionário desenvolvido por um vasto número de cientistas sociais, médicos, enfermeiros e farmacêuticos, cujo objetivo foi criar instrumentos de medição de adesão, fáceis de usar por uma ampla gama de doentes, médicos e investigadores. Foi baseado em pesquisas anteriores sobre a adesão a regimes terapêuticos e ensaios clínicos (Haynes *et al.* (1979), Urquhart (1991), Besch (1995)). O foco é a adesão recente (ou seja, tendo início no dia anterior à data de avaliação) de forma a maximizar a capacidade de recordar e minimizar a existência de viés. Para além disso, foram também avaliadas variáveis que já haviam sido reconhecidas como possíveis determinantes da adesão aos tratamentos tanto da infeção pelo VIH como de outras doenças. São exemplo destas variáveis, o consumo de álcool e drogas, perturbações psicológicas e crenças de auto-eficácia para a adesão (Chesney, 2000).

Os questionários dividem-se em duas versões: o Questionário de Adesão à Medicação ACGT - Versão Base, e o Questionário de Adesão à Medicação ACTG - *Follow-up*, permitindo proceder a uma aplicação única (utilizando os mesmos em conjunto), ou várias aplicações ao longo do tempo (utilizando nas avaliações seguintes apenas o questionário de seguimento (*Follow-Up*)).

Quanto à avaliação das variáveis que se associam com a adesão, as questões agrupam-se em quatro áreas:

Tabela 1.3: Descrição dos métodos indiretos para medição da adesão à terapêutica

Métodos indiretos	
Método	Descrição
Auto-relato da adesão	Acenta na informação prestada pelo doente acerca do seu cumprimento do regime terapêutico. Consiste em solicitar ao indivíduo que informe o profissional de saúde acerca da adesão ao tratamento, normalmente por entrevista, questionário ou preenchimento de um diário em papel ou suporte informático. Tende a sobrevalorizar a adesão, uma vez que o indivíduo poder-se-á sentir intimidado em admitir um eventual não cumprimento do plano terapêutico. Desta forma, a auto-avaliação de uma adesão subótima é um forte indicador de não adesão. Tendo em conta o baixo custo associado, este método é largamente utilizado em estudos e ensaios clínicos sobre a adesão dos doentes à terapêutica adotada para o tratamento de determinada doença
Avaliação clínica	Baseia-se na informação prestada pelo médico e/ou profissionais de saúde sendo, provavelmente, o instrumento mais usual para quantificar a adesão ao tratamento. Apesar disso, apresenta desvantagens, uma vez que um juízo clínico correto está fortemente associado ao conhecimento do doente e das duas características. Perante isto, a avaliação do médico tende a estar baseada em medidas mais objetivas, tais como, a assiduidade às consultas e resultados clínicos
Contagem de forma farmacêuticas sólidas (FFS)	São contados os comprimidos, ou cápsulas, que o doente tem no início e no fim do período em estudo, calculando-se a respetiva diferença e determinando-se a razão com os medicamentos prescritos, para o período estipulado. É um método simples e objetivo, contudo, à semelhança de outros métodos já atrás referidos, pode conduzir a uma sobrevalorização da adesão, uma vez que o facto de faltarem comprimidos não garante que os mesmos tenham sido tomados. Isto acontece porque o doente, quando supõe que se vai proceder à contagem de comprimidos pode retirar os comprimidos da embalagem (denominado de <i>pill-dumping</i>)
Registo da dispensa de medicamentos por parte dos serviços de saúde	Apoia-se na dispensa dos fármacos nos serviços farmacêuticos e é facilmente aplicável. Este instrumento parte do princípio, por vezes, errado, de que o doente não será cumpridor caso não levante a medicação na farmácia, nos intervalos de tempo adequados, os medicamentos necessários para o tratamento. À semelhança das restantes metodologias, o registo da dispensa de medicamentos tem uma limitação óbvia: o facto do indivíduo levantar a medicação não implica que a mesma seja tomada e o esquema terapêutico seja cumprido
Sistemas de monitorização eletrónica	São utilizados desde da década de 80, sendo constituídos por um frasco com uma tampa na qual se encontra inserido um sistema eletrónico que regista o dia e a hora em que o frasco é aberto e fechado. A utilização deste instrumento é considerada uma referência nesta área. O sistema mais conhecido é o <i>Medication Event Monitoring System (MEMS)</i> . Porém, trata-se de um método dispendioso, tendo em conta o elevado preço dos equipamentos e as necessidades logísticas, o que torna impraticável a sua utilização em grandes grupos de indivíduos e por um longo período de tempo

1. **Crenças acerca da medicação** - incluem a auto-eficácia relacionada com a adesão e crenças acerca da eficácia da medicação, utilizando, para tal, uma escala de quatro pontos (varia de "não tem qualquer certeza" a "tem a certeza absoluta"). A expectativa é a de que os doentes que sentem maior auto-eficácia e que pensam que a medicação tem um efeito positivo na saúde, são mais aderentes à medicação;

2. **Perturbações psicológicas e suporte social** - incluem características psicopatológicas, como a depressão, ansiedade e stress, bem como a satisfação com o suporte social percebido (satisfação com amigos, família, atividades sociais e intimidade);
3. **Utilização de álcool e drogas** - Frequência de consumo de bebidas alcoólicas e droga, no mês anterior ao da avaliação;
4. **Características socio-demográficas** - incluem idade, sexo, nível educacional, actividade profissional, existência de crianças a cargo e comportamentos de risco relacionados com o VIH

Relativamente à análise de resultados, é importante salientar que os autores utilizam a resposta à questão "Quando foi a última vez em que falhou alguma toma da sua medicação?" para classificar os doentes como sendo não-aderentes caso refram ter falhado alguma dose durante as duas últimas semanas.

1.2.2 Os fatores associados

De acordo com a OMS (2003), são vários os fatores que afetam a adesão ao regime terapêutico, os quais podem, sumariamente, ser agrupados em cinco grupos:

- **Fatores socio-económicos e culturais:** nível de escolaridade, situação profissional, apoios sociais, condições habitacionais (englobando o número de co-habitantes), preço relacionado com transportes e fármacos, étnia, crenças culturais e desigualdades sociais (Bugalho e Carneiro, 2004). Apesar de, segundo a OMS (2003), os fatores socio-demográficos não estarem claramente associados à adesão, Machado (2009) inclui, neste grupo, alguns destes fatores como a idade, sexo e estado civil;
- **Fatores associados aos profissionais de saúde:** nível de desenvolvimento dos sistemas de saúde, sistema de distribuição de fármacos, taxa de participação, acesso aos medicamentos, recursos humanos e técnicos disponíveis, horários das consultas e conhecimento dos profissionais de saúde sobre doenças crónicas.
- **Fatores relacionados com a doença e co-morbilidade(s):** gravidade dos sintomas, incapacidade física, psicológica, social e profissional, gravidade atribuída à doença por parte do indivíduo e o, consequente, impacto que esta representa na sua vida (Bugalho e Carneiro, 2004). Dentro deste grupo também se engloba o carácter da doença, isto é, se se trata de uma doença aguda ou crónica e se a mesma apresenta ou não sintomas (Giorgi, 2006);

- **Fatores relacionados com a terapêutica:** complexidade (esquemas terapêuticos), duração do tratamento, ausência imediata de melhorias e efeitos secundários (Bugalho e Carneiro, 2004; Giorgi, 2006; Machado, 2009);
- **Fatores relacionados com o doente:** estado psicológico, conhecimentos sobre doença e/ou tratamento, crenças, perceções sobre episódios da doença e as expetativas sobre a evolução da doença. Neste grupo de fatores inclui-se também a falta de capacidade, de motivação e auto-eficácia para gerir o plano terapêutico (Bugalho e Carneiro, 2004; Giorgi, 2006; Oliveira *et al.*, 2007, Machado, 2009).

Apesar desta classificação proposta pela OMS, vários investigadores desta área propõem outras classificações. Almeida *et al.* (2007), por exemplo, indica que os fatores referidos devem ser agrupados em: fatores externos, fatores relacionais e fatores internos ao doente. Os primeiros referem-se ao acesso a medicamentos, às características da doença e ao regime terapêutico. No que concerne aos fatores relacionais, estes autores incluem os apoios sociais e a relação entre o médico e o doente. Por fim, nos fatores internos, são incluídos fatores psicológicos, crenças relativas à saúde e características socio-demográficas.

De entre os fatores externos, estudos demonstram que, quanto maior for o número de medicamentos e/ou tratamentos, menor será a adesão (Leite & Vasconcellos, 2003; Valle, 2000; Santos, 2008). Vários estudos, revelaram, igualmente, que, quanto maior forem os apoios dados ao doente e quanto melhor for a relação médico-doente, melhores são os resultados relativos à adesão (Sabaté, 2003; Leite & Vasconcellos, 2003). Em relação aos fatores internos, não existe unanimidade sobre a influência da idade na adesão. Relativamente ao sexo, este não parece influenciar a adesão, contudo, o baixo nível económico e a baixa escolaridade, parecem influenciar negativamente a adesão (OMS, 2003).

Capítulo 2

Objetivos

O presente estudo tem como objetivo principal a avaliação da prevalência da não-adesão à TARc e a identificação dos fatores (socio-demográficos, biomédicos e psicossociais) associados ao controlo da infecção VIH e à não-adesão à terapêutica anti-retrovírica, permitindo, assim, caracterizar os diferentes perfis dos doentes não aderentes à respetiva terapêutica.

A determinação destes fatores permitirá desenvolver campanhas que incidam sobre os doentes de risco, isto é, indivíduos que se identifiquem com o perfil dos doentes não aderentes, com o objetivo de diminuir a prevalência da não-adesão.

Capítulo 3

Material e Métodos

3.1 Descrição Geral do Estudo

O presente trabalho integra-se no estudo *Adesão à Terapêutica Anti-Retroviral em doentes seropositivos para o VIH: prevalência e fatores associados* (Estudo ATAR-VIH), conduzido pelo Instituto de Medicina Preventiva (IMP) do Hospital de Santa Maria (HSM).

Trata-se de um estudo observacional de coorte: (1) com inclusão retrospectiva de doentes que, entre 2005 e 2008, tenham tido, pelo menos, uma dispensa de TAR, prescrita no Hospital de Dia de Infeciologia do HSM; (2) e com seguimento prospetivo dos doentes sob TAR, durante um período mínimo de 3 meses.

Para o estudo observacional de coorte prospetiva, foram considerados elegíveis todos os indivíduos que verificavam as seguintes condições:

- 1) Ser seropositivo (a) para o VIH-1;
- 2) Maior de 18 anos;
- 3) Seguido(a) no Serviço de Doenças Infeciosas do Hospital de Santa Maria;
- 4) Ter, pelo menos, uma consulta com prescrição de TAR durante o período de recrutamento;
- 5) Capaz de consentir e participar no estudo;
- 6) Não se encontrar detido(a) num estabelecimento prisional, em instituições sociais ou depender de terceiros para a toma da medicação;
- 7) Não incluídos em ensaios clínicos.

No período de abril a julho de 2011, foi pedido aos médicos assistentes que identificassem os indivíduos elegíveis de forma sistemática, no momento da consulta, e que os convidassem a participar no estudo. Após consentimento informado, foi aplicado um questionário aos participantes em entrevista presencial e pedido aos médicos o preenchimento de um formulário com dados clínicos. A adesão à terapêutica foi avaliada com base no Questionário AIDS Adult Clinical Trial Group (AACTG), desenvolvido para auto-avaliação da adesão recente (últimos 4 dias, último fim-de-semana e os últimos 30 dias). Desta forma, dos métodos referidos no capítulo anterior, foi utilizado um método indireto de avaliar a adesão. Em particular, o auto-relato da adesão.

Em função das respostas para os diferentes momentos avaliados, foi possível classificar cada indivíduo como apresentando uma adesão elevada, moderada ou fraca (figura 3.1).

Tabela 3.1: Classificação dos níveis de adesão à terapêutica anti-retrovírica, em função das respostas dadas para os diferentes momentos ou períodos de tempo avaliados

Momento	Adesão elevada	Adesão moderada	Adesão fraca
Últimos 4 dias	Sempre	Sempre	Falhou uma toma
Posologia nos últimos 4 dias	Sempre	Frequentemente ou menos	— — —
Fim de semana	Cumpriu	Não cumpriu	— — —
Últimos 30 dias	Sempre	Ocasional ou frequentemente	Raramente ou nunca

Posteriormente, classificaram-se como aderentes os participantes considerados com adesão elevada e como não aderentes, aqueles que foram cuja adesão à terapêutica foi considerada moderada ou fraca, de acordo com a classificação apresentada na tabela 3.1.

O questionário ao participante incluiu, para além do Questionário AACTG, a avaliação de características sócio-demográficas e clínicas e a aplicação de escalas validadas para a população portuguesa, para avaliação de a) Satisfação com o Suporte Social (ESSS), cuja pontuação varia 15 e 75 pontos - quanto maior a pontuação obtida, maior a satisfação com o suporte social; b) Ansiedade, Depressão e Stress (EADS), cuja pontuação oscila entre 0 entre 21 pontos - quanto maior a pontuação, mais negativo é o estado psicológico; c) Crenças e Perceção da Doença (IPQ-R), cuja pontuação varia 0 entre 45 pontos - quanto maior a pontuação mais crónica é a perceção da doença; d) Crenças sobre a medicação (BMQ), onde quanto maior a pontuação, maiores são as crenças acerca da necessidade dos fármacos e das crenças acerca dos efeitos secundários do mesmo.

3.1.1 Escala de Satisfação com o Suporte Social (ESSS)

A **Escala de Satisfação do Suporte Social (ESSS)** é um instrumento desenvolvido por Wethington and Kessler (1986) e validado para a população portuguesa por Pais-Ribeiro (1999).

A ESSS é um questionário de auto-preenchimento constituído por 15 itens sob a forma de afirmações. O grau com que o indivíduo concorda com as afirmações é medido através de uma escala de tipo Likert com cinco opções de resposta: "discordo plenamente", "discordo", "não concordo nem discordo", "concordo" e "concordo plenamente".

Esta escala é composta por quatro subescalas, cujo objetivo é o de avaliar diferentes parâmetros do suporte social:

- **Satisfação com amigos (SA)** - constituída por cinco itens (3, 12, 13, 14 e 15), mede a satisfação com as amizades que o indivíduo possui. Nesta dimensão são colocadas várias afirmações, tais como:

"Os amigos não me procuram tantas vezes quanto gostaria."

"Estou satisfeito com a quantidade de amigos que tenho."

- **Intimidade (IN)** - constituída por quatro itens (1, 4, 5 e 6), avalia a existência de suporte de apoio social íntimo. São exemplos das afirmações:

"Por vezes sinto-me só no mundo se sem apoio."

"Quando preciso de desabafar com alguém encontro facilmente amigos com quem o fazer."

- **Satisfação com a família (SF)** - Constituída por três itens (9, 10 e 11), mede a satisfação com o suporte familiar existente. Como exemplo, temos as seguintes afirmações:

"Estou satisfeito com a quantidade de tempo que passo com a minha família."

"Estou satisfeito com o que faço em conjunto com a minha família."

- **Atividades sociais (AS)** - Constituída por três itens (2, 7 e 8), avalia a satisfação com as atividades sociais desenvolvidas. Nesta dimensão é avaliado o grau de concordância com várias afirmações, das quais, de seguida, se indicam dois exemplos:

"Não saio com amigos tantas vezes quantas eu gostaria."

"Sinto falta de atividades sociais que me satisfaçam."

Os itens são cotados de 1 a 5 de acordo com a grelha de correção que se encontra na tabela B.1.

A classificação de cada subescala é resultado da soma dos itens que a compõem e a soma da pontuação das quatro subescalas dá a nota total da escala. Desta forma, a ESSS varia entre 15 e 75.

Quanto maior é o valor total da escala, maior é o suporte social percebido.

A escala validada por Pais-Ribeiro (1999) tem uma consistência interna (Alfa de Cronbach) de 0,85. No que respeita às subescalas, a consistência interna obtida encontra-se na tabela B.1.

Tabela 3.2: Alfa de Cronbach para as subescalas do ESSS

Subescala	Alfa de Cronbach
Satisfação com amigos	0,83
Intimidade	0,74
Satisfação com a família	0,74
Atividades sociais	0,64

3.1.2 Escala de Ansiedade, Depressão e Stress (EADS)

A escala de Ansiedade, Depressão e Stress, originalmente, proposta por Lovibond and Lovibond (1995), mais tarde, foi validada para a população portuguesa por Pais-Ribeiro *et al.* (2004), denominada por Escala de Ansiedade, Depressão e Stress (EADS).

A EADS é constituída por 21 itens, os quais se distribuem por três dimensões com sete itens cada: Depressão, Ansiedade e Stress. A escala foi desenvolvida através da teoria existente, pelo que os itens incluídos em cada uma das dimensões propõem avaliar os aspetos relacionados com a respetiva dimensão. Nomeadamente, a depressão abrange: disforia (um item); desânimo (um item); desvalorização da vida (um item), auto-depreciação (um item); falta de interesse ou envolvimento (um item); anedonia (um item); inércia (um item). A Ansiedade engloba: excitação do sistema autónomo (três itens); efeitos músculo-esqueléticos (um item); ansiedade situacional (um item); experiências subjetivas de ansiedade (dois itens). Por último, o stress é composto por: dificuldade em relaxar (dois itens); excitação nervosa (um item); facilmente agitado/chateado (um item); irritável/reação exagerada (dois itens); impaciência (um item) (Pais-Ribeiro *et al.*, 2004).

A resposta é dada numa escala tipo Likert, em que o indivíduo avalia a extensão em que experimentou cada sintoma durante a última semana (últimos 7 dias), numa escala de quatro pontos "Não se aplicou nada a mim", "Aplicou-se a mim algumas vezes", "Aplicou-

se a mim muitas vezes" e "Aplicou-se a mim a maior parte das vezes", a que correspondem, respetivamente, os valores de 0 a 3, conforme consta na tabela B.2.

A EADS fornece três pontuações, uma para cada dimensão, sendo que a pontuação mínima é 0 e a pontuação máxima é 21. A pontuação de cada uma das subescalas é resultado da soma da pontuação dos 7 itens respetivos. Os itens da EADS de 21 itens foram selecionados de modo a que possa ser convertida na escala completa de 42 itens multiplicando a nota por dois.

Para cada uma destas escalas foram utilizados os seguintes construtos:

- **Depressão (D)** - Disforia (item 13); desânimo (item 10); desvalorização da vida (item 21); auto-depreciação (item 17); falta de interesse ou de envolvimento (item 16), anedonia (item 3) e inércia (item 5).

Como exemplo das afirmações relacionadas com esta dimensão tem-se:

"Não consegui sentir nenhum sentimento positivo." (anedonia)

"Tive dificuldade em tomar iniciativa para fazer coisas." (inércia)

- **Ansiedade (A)** - Excitação do sistema autónomo (itens 2, 4 e 19); efeitos músculo-esqueléticos (item 7); ansiedade situacional (item 9); experiências subjetivas de ansiedade (itens 15 e 20).

São exemplos, as seguintes afirmações:

"Senti a minha boca seca." (excitação do sistema autónomo)

"Senti tremores (por ex., nas mãos)." (efeitos músculo-esqueléticos)

- **Stress (S)** - Dificuldade em relaxar (itens 1 e 12); excitação nervosa (item 8); facilmente agitado/chateado (item 18); irritável/reação exagerada (itens 6 e 11) e impaciência (item 14).

Como exemplo, temos as seguintes afirmações:

"Tive dificuldades em me acalmar." (dificuldade em relaxar)

"Estive intolerante em relação a qualquer coisa que me impedisse de terminar aquilo que estava a fazer." (impaciência)

Uma vez que cada uma das subescalas possui sete itens, a pontuação das mesmas varia no intervalo [0,7], sendo que as notas mais elevadas correspondem estados afetivos mais negativos.

No que respeita à consistência interna deste instrumento, Pais-Ribeiro *et al.* (2004) verificaram que este possuía boas propriedades psicométricas, tendo em conta o Alfa de Cronbach determinados e os quais se encontram na tabela 3.3.

Tabela 3.3: Alfa de Cronbach para as subescalas do EADS

Subescala	Alfa de Cronbach
Depressão 7 itens (14 itens)	0,85 (0,93)
Ansiedade 7 itens(14 itens)	0,74(0,83)
Stress 7 itens (14 itens)	0,81 (0,88)

3.1.3 Questionário da Perceção da Doença

O questionário da perceção da doença, validado para a população portuguesa por Figueiras *et al.* (2002) foi originalmente, concebido por Moss-Morris *et al.* (2002), denominado de *Revised Illness Perception Questionnaire (IPQ-R)*.

O questionário possui como pressuposto básico o facto de os doentes criarem os próprios modelos ou representações relativos às suas doenças por forma a dar sentido e resolver o problema com que se deparamm adaptando como moldura teórica o modelo de auto-regulação de Leventhal.

O *IPQ-R* é dividido em três secções, com a identidade e as dimensões causais apresentadas separadamente das restantes, uma vez que são específicas de cada doença e mudam consoante a patologia em estudo. Na versão revista (utilizada no presente estudo), são considerados catorze sintomas. É solicitado aos participantes que indiquem se apresentam esse sintoma desde o início da doença, através de uma resposta sob a forma de sim/não. De seguida, devem indicar se acreditam que o mesmo está relacionado com a doença, através do mesmo formato de resposta.

Posteriormente, são apresentados 38 itens distribuídos em 7 subescalas: dimensão temporal aguda/crónica (6 itens), dimensão cíclica (carácter intermitente de sinais e/ou sintomas da doença) (4 itens), consequências (6 itens), controlo pessoal (6 itens), controlo de tratamento (5 itens), coerência da doença (5 itens) e representação emocional (6 itens). Estes itens são avaliados na escala de tipo Likert de 5 pontos: "discorda plenamente", "discorda", "não concorda nem discorda", "concorda", "concorda plenamente". A cada um dos itens é atribuída uma cotação em função da resposta facultada, de acordo com a grelha de correção que se encontra nas tabelas B.3 e B.4.

Como exemplo as afirmações para cada uma das dimensões que compõem o *IPQ-R* tem-se:

- Duração aguda/crónica:

"A minha doença dura pouco tempo."

"A minha doença parece ser mais permanente que temporária."

- Duração cíclica:

"Os sintomas da minha doença vão e vêm em ciclos."

"A minha doença é imprevisível."

- Consequências:

"A minha doença é grave."

"A minha doença tem grande consequências para a minha vida."

- Controlo pessoal:

"Há muita coisa que eu posso fazer para controlar os meus sintomas."

"O que eu faço pode determinar se a minha doença melhora ou piora."

- Controlo de tratamento:

"O tratamento é eficaz na cura da minha doença."

"O tratamento pode controlar a minha doença."

- Coerência da doença:

"Os sintomas da minha doença confundem-me."

"A minha doença é um mistério para mim."

- Representação emocional:

"Fico deprimido(a) quando penso sobre a minha doença."

"A minha doença faz-me sentir zangado(a)."

Tendo em conta o número de itens que compõem cada uma das subescalas, tem-se que o intervalo de variação de cada subescala é: identidade (0-14), duração aguda/crónica (6-30), duração cíclica (4-20), consequências (6-30), controlo pessoal (6-30), controlo de tratamento (5-25), coerência da doença (5-25) e representação emocional (6-30).

Valores elevados nas escalas de identidade, duração aguda/crónica, consequências, e duração cíclica refletem ideias negativas sobre o número de sintomas, a duração crónica da doença, as consequências da doença e a natureza cíclica da enfermidade, respetivamente. Por outro lado, pontuações elevadas para as dimensões controlo pessoal, controlo de tratamento e coerência da doença, representam crenças positivas sobre estas 3 subescalas (Moss-Morris *et al.* (2002); Figueiras *et al.* (2002); Santos and Lopes (2003)).

Este instrumento foi validado numa amostra composta por 711 indivíduos, pertencentes a oito grupos de doenças: asma (86), diabetes(73), artrite reumatóide (76), dor crónica (63), dor aguda (35), enfarte de miocárdio (47), esclerose múltipla (170) e VIH (161). No que concerne às propriedades psicométricas, todas as subescalas revelaram boa consistência interna, como se pode verificar na tabela 3.4.

Tabela 3.4: Alfa de Cronbach para as subescalas do *IPQ-R*

Subescala	Alfa de Cronbach
Duração aguda/crónica	0,89
Duração cíclica	0,79
Consequências	0,84
Controlo pessoal	0,81
Controlo de tratamento	0,80
Coerência da doença	0,87
Representação Emocional	0,88

O *IPQ-R* possui a vantagem de poder ser adaptado consoante a doença que se pretende estudar, substituindo a palavra doença pelo nome específico da respetiva enfermidade, como por exemplo, VIH, diabetes, entre outros. Para além disso, itens importantes para as subescalas identidade e causa da doença podem ser facilmente inseridas, tendo em conta a estrutura do instrumento em três secções.

No endereço <http://www.uib.no/ipq>, encontram-se, à semelhança do que acontece para o *IPQ*, várias traduções do instrumento *IPQ-R*, em particular, a versão original, na língua inglesa (Moss-Morris *et al.*, 2002) e a versão em português, validada para a população portuguesa por Figueiras *et al.* (2002). Está igualmente disponível a versão inglesa para a doença crónica VIH, entre outras doenças.

3.1.4 Questionário de Crenças Acerca da Medicação

Horne *et al.* (1999) desenvolveram o Questionário de Crenças Acerca da Medicação (*Beliefs about Medicine Questionnaire - BMQ*), cujo objetivo é avaliar as representações cognitivas que os indivíduos fazem em relação aos fármacos, considerando esta área como um desafio significativo nos domínios da investigação e da prática na área da saúde. Segundo estes autores, O desenvolvimento desta escala justificou-se, segundo os autores pelo facto de a prescrição de fármacos preencher uma vasta gama das intervenções médicas e constituir uma das principais fontes de despesas na área da saúde na maioria dos países industrializados por um lado, e pelos conhecidos índices de não-adesão a estas prescrições por outro.

A versão original é composta por duas secções, designadas por: (1) BMQ-específica, e (2) BMQ-geral. A primeira avalia as representações de fármacos prescritos para uso pessoal, enquanto a segunda avalia crenças acerca dos medicamentos em geral.

A BMQ-específica, utilizada no presente estudo, é composta por dois fatores de cinco itens que avaliam as crenças dos indivíduos acerca da **necessidade** da medicação que lhes foi prescrita, e as **preocupações** relacionadas com a mesma com base em crenças acerca dos perigos de dependência e toxicidade ou efeitos secundários a longo prazo.

A BMQ-geral é composta por dois fatores de quatro itens que avaliam as crenças relacionadas com potenciais efeitos nocivos, aditivos e tóxicos dos medicamentos (*harm*), e com a possibilidade de estes serem usados em excesso pelos médicos, isto é, uso excessivo (*overuse*).

Estas duas secções do BMQ podem ser usadas em combinação ou separadamente consoante os objetivos do estudo.

O grau com que o indivíduo concorda com as afirmações é medido através de uma escala de tipo Likert com cinco opções de resposta: "discordo plenamente", "discordo", "não concordo nem discordo", "concordo" e "concordo plenamente".

O questionário é subdividido em duas escalas ou fatores:

- **Preocupações**, com onze itens (itens: 2, 5, 6, 8, 9, 11, 13, 17, 18, 19, e 21).

Como exemplo, tem-se as seguintes afirmações:

"Preocupa-me ter que tomar esta medicação."

"Por vezes, preocupo-me com os efeitos a longo prazo desta medicação."

- **Necessidades**, com oito itens (itens: 1, 3, 4, 7, 10, 14, 15 e 20).

São exemplos, as seguintes afirmações:

"Atualmente, a minha saúde depende da medicação anti-retrovírica que me foi prescrita."

"A minha vida seria impossível sem esta medicação."

Após a cotação inversa dos itens 14, 17 e 19 (tabela B.5), a classificação de cada subescala é resultado da soma dos itens que a compõem. Os itens 12 e 16 foram recentemente retirados da escala pelos autores, dado que se referem a dificuldades práticas na toma da medicação, pelo que não procedemos, igualmente, à sua análise.

Na BMQ-específica, quanto maior a pontuação obtida em cada uma das duas subescalas, maiores são as crenças na necessidade da toma de medicação e as crenças acerca dos efeitos secundários dos medicamentos.

3.1.5 Variáveis em estudo

Na figura 3.1 apresentam-se as variáveis previstas no modelo concetual, consideradas como potenciais fatores de não-adesão.

Doente	Situação clínica	Crenças sobre a doença e os fármacos
<ul style="list-style-type: none"> • Idade • Género • País de nascimento • Etnia • Estado Civil • Filhos menores • Mora sozinho • Número de co-habitantes • Nível de escolaridade • Situação profissional • Horário de trabalho • Prática de exercício físico • Hábitos alcoólicos • <i>Binge-drinking</i> • Hábitos tabágicos • Outros consumos • Tratamento sob metadona • Satisfação com o suporte social • Ansiedade • Depressão • Stress • Via de transmissão 	<ul style="list-style-type: none"> • Classificação SIDA • Presença de co-infecções • Presença de co-morbilidades • Situação clínica na 1ª consulta (assintomático, sintomático ou diagnóstico SIDA) • Situação clínica aquando início da TAR (assintomático, sintomático ou diagnóstico SIDA) • Situação clínica no momento da aplicação do questionário (assintomático, sintomático ou diagnóstico SIDA) • Classificação CDC na 1ª consulta • Classificação CDC aquando início da TAR • Classificação CDC no momento da aplicação do questionário • Efeitos adversos • Adesão estimada pelo médico • Recomendação sobre TAR • Recomendação de adesão • Mudança de regime • Tempo sob TAR 	<ul style="list-style-type: none"> • Crenças acerca da doença: • Identidade • Duração aguda/crónica • Duração cíclica • Consequências • Controlo pessoal • Controlo de tratamento • Coerência da doença • Representação emocional • Crenças sobre a medicação: • Preocupações • Necessidades

Figura 3.1: Variáveis em estudo.

3.2 Análise Estatística

Para a caracterização da amostra foi utilizada a estatística descritiva, no que diz respeito às variáveis relacionadas com o doente, situação clínica e a relação com os serviços de saúde.

A análise estatística foi realizada no software R, versão 3.0.1 (www.r-project.org).

Com base nas variáveis descritas como potenciais fatores da não-adesão (figura 3.1) e assumindo a não-adesão como uma variável dicotómica, comparou-se o grupo dos aderentes com o dos não aderentes, com recurso ao teste do Qui-quadrado (χ^2) (ou teste exato de Fisher, quando não se encontravam reunidas as condições de aplicação do teste do χ^2)

para as variáveis categóricas. Para as variáveis contínuas, a comparação entre os dois grupos foi efetuada através do teste de Mann-Whitney.

Para analisar a influência conjunta das variáveis de interesse na discriminação do perfil dos aderentes e não aderentes utilizou-se a análise de regressão logística. Inicialmente, foi efetuada a análise preliminar univariada de forma a determinar as variáveis associadas mais significativamente à variável resposta ($\text{valor-}p \leq 0,15$). Em seguida, estas variáveis foram submetidas à análise multivariada através do modelo de regressão logística com a finalidade de identificar os preditores de não-adesão. Foi utilizado o critério AIC (*Akaike's an information criterion*) para a seleção do melhor modelo, que tem como base o teste de razão de verossimilhanças, o qual consiste na comparação estatística, por χ^2 , dos resíduos dos desvios, do modelo com melhor ajuste em relação ao restantes. A variável foi excluída se o valor-p para o teste de razão de verossimilhanças for inferior a 0.05. Foi utilizado um método stepwise (*stepwise forward*) para a seleção das variáveis a incluir no modelo ($p \leq 0,05$). Foram ainda calculadas as estimativas do *odds ratio* (OR) e dos intervalos de confiança (IC) a 95%, para as variáveis não excluídas.

Com o objetivo de avaliar a qualidade do modelo, foi realizado o teste de Hosmer-Lemeshow e determinada a área sob a curva ROC (*Receiver Operating Characteristic*) e respetivas sensibilidade e especificidade.

No que respeita à metodologia de Análise de Correspondências Múltiplas, numa primeira abordagem, foram consideradas as variáveis submetidas ao modelo de regressão logística múltipla, de forma a verificar as relações mais fortes e identificar, assim, os diferentes perfis de doentes. Por fim, consideraram-se apenas as variáveis que entraram no modelo final de regressão logística múltipla e sobre estas foi aplicada novamente a metodologia via ACM.

Capítulo 4

Modelo Linear Generalizado

4.1 Introdução

Durante um longo período de tempo, os modelos normais lineares foram utilizados na tentativa de descrever a maioria dos fenômenos aleatórios. Mesmo quando o fenômeno sob estudo não apresentava uma resposta para a qual fosse plausível assumir a normalidade, era sugerido algum tipo de transformação com a finalidade de alcançar a tão ambicionada normalidade. A mais conhecida, porventura, foi proposta por Box e Cox (1964), a qual transforma o valor observado y (positivo) em

$$z = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(y), & \lambda = 0 \end{cases}$$

onde λ é uma constante desconhecida.

Com o desenvolvimento computacional ocorrido na década de 70, tornou-se frequente a utilização de processos iterativos para a estimação dos parâmetros. Porém, a proposta mais aliciente e inovadora foi apresentada por Nelder e Wedderburn (1972), que propuseram os **Modelos Lineares Generalizados** ou **GLM** (*Generalized Linear Models*), classe muito vasta da qual fazem parte os modelos lineares.

Estes modelos surgiram na tentativa de descrever situações nas quais a variável resposta não segue uma distribuição Gaussiana, não se podendo utilizar:

- i) O modelo linear (ML), no caso em que as observações são independentes;
- ii) O modelo linear misto (MLM), quando as observações não são independentes.

Assim, os Modelos Lineares Generalizados, doravante designados por MLG, correspondem a uma síntese dos modelos lineares (ML) e de outros modelos, tendo sido unificada, tanto do ponto de vista teórico como concetual, a teoria da modelação estatística até então desenvolvida.

O crescente interesse pela área conduziu à realização de vários encontros informais no início dos anos 80, a maioria deles em Inglaterra, até que, em 1986, foi realizado na cidade de Innsbruck (Áustria), o "*1st International Workshop on Statistical Modelling*" (*1st IWSM*). Este encontro tem sido realizado anualmente, sendo que o último (28th IWSM) realizou-se entre 8 e 12 de julho de 2013, em Palermo (Itália).

Os MLG são atualmente uma vasta classe de modelos de regressão que incluem, em particular, e para citar os mais comuns na área biomédica, os modelos de regressão linear clássicos, de análise de variância (ANOVA) com distribuição normal e variável resposta contínua, de regressão logística com variável resposta binária ou dicotómica e de regressão de Poisson ou log-linear para contagens.

A extensão do modelo linear é feita em duas direções:

1. A variável resposta pode ter uma distribuição que não seja normal;
2. A relação entre a variável resposta e as variáveis explicativas não necessita de ser linear.

Para que seja possível aplicar o MLG, é necessário que a variável resposta possua distribuição pertencente à família exponencial, pelo que, antes de mais, se irá apresentar a definição de *Família Exponencial*.

Definição 4.1 (Família Exponencial).

Diz-se que uma variável aleatória Y tem distribuição pertencente à família exponencial de dispersão (ou simplesmente, família exponencial) se a sua função densidade de probabilidade (f.d.p.) ou função massa de probabilidade (f.m.p.) se puder escrever na forma

$$f(y|\theta, \phi) = \exp \left\{ \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi) \right\}, \quad (4.1)$$

onde θ e ϕ são parâmetros escalares (localização e dispersão, respetivamente), $a(\cdot)$, $b(\cdot)$ e $c(\cdot, \cdot)$ são funções reais conhecidas.

Na definição (4.1), θ é a **forma canónica do parâmetro de localização** e ϕ é o **parâmetro de dispersão ou escala** suposto, em geral, conhecido. Por vezes, este parâmetro é denotado por σ^2 . Neste caso, a distribuição enunciada em (4.1) faz parte

da família exponencial univariada. Quando o parâmetro ϕ é desconhecido, a distribuição pode ou não fazer parte da família exponencial biparamétrica, tal como é geralmente definida. Este parâmetro é constante ao longo das observações. Admite-se ainda que $b(\cdot)$ é diferenciável e que o suporte da distribuição não depende dos parâmetros.

Nestas circunstâncias, a família exponencial obedece às condições de regularidade.

Em muitas situações de interesse, observa-se que a função $a(\phi)$ toma a forma $a(\phi) = \frac{\phi}{\omega}$, onde ω é uma constante conhecida, obtendo-se a variância de Y como o produto do parâmetro de dispersão por uma função apenas do valor médio. Neste caso, a função definida em (4.1) escreve-se na forma

$$f(y|\theta, \phi, \omega) = \exp \left\{ \frac{\omega}{\phi} (y\theta - b(\theta)) + c(y, \phi, \omega) \right\}, \quad (4.2)$$

onde ω é uma constante conhecida e que varia de observação para observação e à qual se dá o nome de **peso**.

4.1.1 Valor Médio e Variância

Seja $l(\theta; \phi, y) = \ln(f(y|\theta, \phi))$, isto é, l é o logaritmo neperiano da f.d.p. ou f.m.p. de Y .

A função **score**¹ é dada por

$$S(\theta) = \frac{\partial l(\theta; \phi, Y)}{\partial \theta} \quad (4.3)$$

Para famílias regulares, tem-se:

$$E(S(\theta)) = 0 \quad (4.4)$$

$$E(S^2(\theta)) = E \left[\left(\frac{\partial l(\theta; \phi, Y)}{\partial \theta} \right)^2 \right] = -E \left[\frac{\partial^2 l(\theta; \phi, Y)}{\partial \theta^2} \right] \quad (4.5)$$

Assim sendo, no caso em que $f(y|\theta, \phi)$ é dado por (4.1),

$$l(\theta; \phi, y) = \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi), \quad (4.6)$$

obtém-se

$$S(\theta) = \frac{Y - b'(\theta)}{a(\phi)} \Rightarrow \frac{\partial S(\theta)}{\partial \theta} = -\frac{b''(\theta)}{a(\phi)}, \quad (4.7)$$

¹denotada, por vezes, por U .

onde $b'(\theta) = \frac{\partial b(\theta)}{\partial \theta}$ e $b''(\theta) = \frac{\partial^2 b(\theta)}{a(\phi)}$. Deste modo, de 4.4, 4.5 e 4.7, sai que

$$E(Y) = \mu = a(\phi)E(S(\theta)) + b'(\theta) = b'(\theta) \quad (4.8)$$

$$\text{var}(Y) = a^2(\phi)\text{var}(S(\theta)) = a^2(\phi)\frac{b''(\theta)}{a(\phi)} = a(\phi)b''(\theta) \quad (4.9)$$

Verifica-se, assim, que a variância de Y depende da função $b''(\theta)$, isto é, depende do valor médio. A esta função é dado o nome de **função de variação** ou de **variância**, a qual é denotada por $V(\mu_i)$, donde $V(\mu_i) = b''(\theta)$.

De referir que a família exponencial inclui várias distribuições, tais como, distribuição Normal (ou Gaussiana), Bernoulli, Binomial, Poisson e Binomial Negativa (a prova não será apresentada uma vez que foge do âmbito do presente estudo).

4.1.2 Extensão do ML ao MLG

Considerem-se n observações independentes da variável resposta Y e seja Y_i , $i = 1, \dots, n$, a variável resposta para o i -ésimo indivíduo. Seja $y = (y_1, \dots, y_n)^T$ o vetor de observações, em que y_i é a observação da variável resposta para o i -ésimo indivíduo. Associado a cada resposta, Y_i , encontra-se o vetor de covariáveis de dimensão $(k \times 1)$, (x_{i1}, \dots, x_{ik}) , com $i = 1, \dots, n$ e onde x_{ij} corresponde à j -ésima covariável para o i -ésimo indivíduo. Seja ainda $\beta = (\beta_1, \dots, \beta_k)^T$ um vetor de dimensão $(k \times 1)$ de parâmetros desconhecidos. Na maioria dos casos, $x_{i1} = 1$, para qualquer i .

A parte sistemática do ML assume a forma

$$\mu_i = \sum_{j=1}^k x_{ij}\beta_j = x_i^T \beta, \quad i = 1, \dots, n, \quad (4.10)$$

onde x_{ij} é o valor da j -ésima covariável para a observação i e β_j o j -ésimo parâmetro desconhecido. Em notação matricial escreve-se na forma

$$\mu = X\beta, \quad (4.11)$$

onde μ é um vetor de dimensão $(n \times 1)$.

Ao definir-se o **preditor linear** como $\eta_i = x_i^T \beta$, a relação entre o valor médio, μ_i , e o preditor linear, η_i , é a identidade. Existe, neste caso, uma relação linear direta entre μ_i e as variáveis independentes.

De forma a simplificar a transição dos ML para os MLG, pode-se especificar o ML em três componentes que são dadas por:

1. **componente aleatória:** as variáveis aleatórias Y_i são independentes e têm distribuição normal com $E(Y_i) = \mu_i$ e a variância, σ^2 , constante, isto é,

$$Y_i \sim N(\mu_i, \sigma^2) \quad i = 1, \dots, n; \quad (4.12)$$

2. **Componente sistemática:** Define-se um preditor linear η_i como combinação das variáveis explicativas

$$\eta_i = x_i^T \beta; \quad (4.13)$$

3. **Função de ligação** entre as componentes aleatória e sistemática:

$$g(\mu_i) = \eta_i. \quad (4.14)$$

Como já fora referido, anteriormente, os MLG são obtidos através da extensão das hipóteses subjacentes aos ML, em duas direções, mais precisamente:

1. A distribuição de Y_i pode ser qualquer desde que pertença à família exponencial, atrás definida;
2. Permite a existência de outras formas de ligação, além da identidade, entre o preditor linear, μ_i , e o valor médio, μ_i , ou seja,

$$\eta_i = g(\mu_i), \quad (4.15)$$

onde $g(\cdot)$ é uma função monótona e diferenciável à qual se dá o nome de **função de ligação**.

Quando se verifica $\eta_i = \theta_i$ dá-se, a $g(\cdot)$, o nome de **função de ligação canónica**.

Esquematicamente, o modelo de Regressão Logística pode ser apresentado pela figura 4.1.

4.2 Regressão Logística

4.2.1 O Modelo

Geralmente, nos estudos epidemiológicos, a grande questão à qual se pretende obter resposta, é saber qual a relação entre uma ou mais variáveis e a doença. Deste modo, a

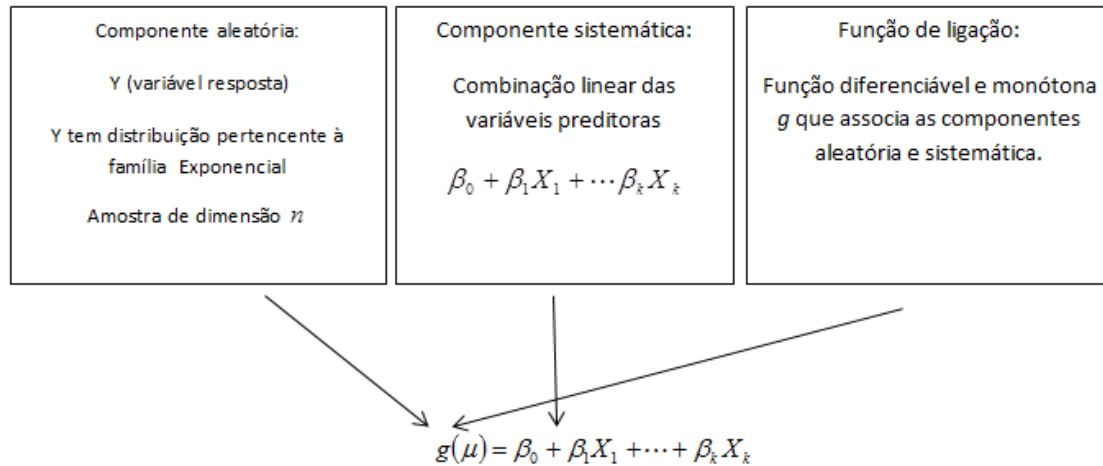


Figura 4.1: Estrutura do modelo de Regressão Logística.

regressão logística representa um dos principais métodos utilizados na área da epidemiologia. Mesmo quando a resposta não é binária, alguns autores dicotomizam a resposta de modo a que a probabilidade de sucesso possa ser ajustada através da regressão logística. Na análise de sobrevivência discreta, por exemplo, na qual a resposta de interesse é o tempo de sobrevivência, em algumas pesquisas, a função de risco tem sido ajustada para modelos logísticos. Tudo isto se deve, principalmente, à facilidade de interpretação dos parâmetros de um modelo logístico e também pela possibilidade de ser utilizada na análise discriminante.

Embora a regressão logística seja conhecida desde os anos 50, foi Cox (1970) que a tornou popular entre o meio estatístico. Em muitas situações, a variável resposta Y assume o valor 1 ou 0 consoante se observa a presença ($Y = 1$) ou ausência ($Y = 0$) de determinada característica em estudo.

O modelo de regressão logística é adequado para dados binomiais ou em forma de proporções.

Veja-se, de seguida que, em ambos os casos, a distribuição subjacente à variável resposta pertence à família exponencial.

Dados binomiais

Seja Y uma variável aleatória com distribuição Binomial de parâmetros m e π , isto é, $Y \sim \text{Bin}(m, \pi)$. A sua f.m.p. é dada por

$$\begin{aligned}
f(y|\pi) &= \binom{m}{y} \pi^y (1-\pi)^{m-y} \\
&= \exp \left\{ \ln \left[\binom{m}{y} \pi^y (1-\pi)^{m-y} \right] \right\} \\
&= \exp \left\{ \ln \binom{m}{y} + \ln(\pi^y) + \ln((1-\pi)^{m-y}) \right\} \\
&= \exp \left\{ \ln \binom{m}{y} + y \ln \pi + (m-y) \ln(1-\pi) \right\} \\
&= \exp \left\{ y \ln \pi + m \ln(1-\pi) - y \ln(1-\pi) + \ln \binom{m}{y} \right\} \\
&= \exp \left\{ y (\ln \pi - \ln(1-\pi)) + m \ln(1-\pi) + \ln \binom{m}{y} \right\} \\
&= \exp \left\{ y \ln \left(\frac{\pi}{1-\pi} \right) + m \ln(1-\pi) + \ln \binom{m}{y} \right\}
\end{aligned}$$

Deste modo, a v.a. Y tem distribuição pertencente à família exponencial, com

$$\theta = \ln \left(\frac{\pi}{1-\pi} \right)$$

$$b(\theta) = -m \ln(1-\pi) = m \ln(1+e^\theta) \quad ^2$$

$$c(y, \phi) = \ln \binom{m}{y}$$

$$a(\phi) = 1, \quad \phi = \omega = 1$$

De notar que

$$b'(\theta) = m \frac{e^\theta}{1+e^\theta} = m \frac{\frac{\pi}{1-\pi}}{1+\frac{\pi}{1-\pi}} = \frac{m(1-\pi)\pi}{1-\pi} = m\pi$$

$$b''(\theta) = m \frac{e^\theta}{(1+e^\theta)^2} = m \frac{\frac{\pi}{1-\pi}}{\left(1+\frac{\pi}{1-\pi}\right)} = m\pi(1-\pi)$$

De (4.8) e (4.9), obtém-se, diretamente,

$$E(Y) = b'(\theta) = m\pi, \quad \text{var}(Y) = b''(\theta)a(\phi) = m\pi(1-\pi)$$

2

$$\theta = \ln \left(\frac{\pi}{1-\pi} \right) \Leftrightarrow e^\theta = \frac{\pi}{1-\pi} \Leftrightarrow e^\theta(1-\pi) = \pi \Leftrightarrow e^\theta = \pi + \pi e^\theta \Leftrightarrow e^\theta = \pi(1+e^\theta) \Leftrightarrow \frac{e^\theta}{1+e^\theta} = \pi.$$

$$\text{Portanto, } -m \ln(1-\pi) = -m \ln \left(1 - \frac{e^\theta}{1+e^\theta} \right) = -m \ln \left(\frac{1+e^\theta - e^\theta}{1+e^\theta} \right) = -m \ln(1+e^\theta)^{-1} = m \ln(1+e^\theta)$$

Assim, considerando as variáveis resposta Y_i , $i = 1, \dots, n$, tal que $Y_i \sim \text{Bin}(m_i, \pi_i)$, com $m_1 = \dots = m_n = 1$, tem-se que $E(Y_i) = \pi_i$ e $\theta_i = \ln\left(\frac{\pi_i}{1 - \pi_i}\right)$, tem-se que, ao fazer-se

$$\theta_i = \eta_i = x_i^T \beta$$

conclui-se que, a função de ligação canônica é a função *logit* (*logistic inverse transformation*).

Na figura 4.2 verifica-se que, efetivamente, a função *logit* verifica as condições de uma função de ligação, isto é, é monótona e diferenciável no intervalo $]0,1[$.

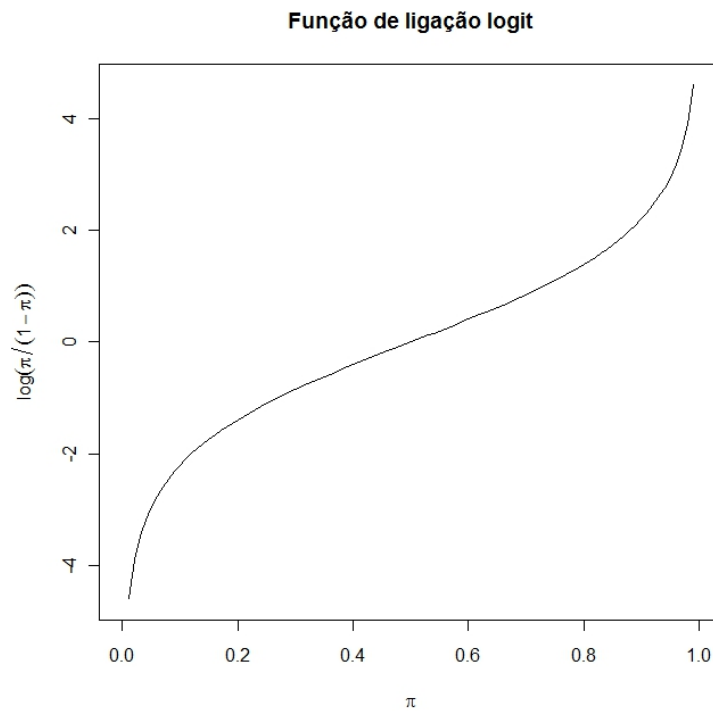


Figura 4.2: Representação gráfica da função *logit*.

Desta forma, a probabilidade de sucesso, i.e., $\mathcal{P}(Y_i = 1) = \pi_i$, está relacionada com o vetor x_i através da seguinte relação

$$\pi_i = \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)} \quad (4.16)$$

Caso existam k variáveis aleatórias independentes (X_1, \dots, X_k) , o modelo é dado por

$$\pi_i = \frac{e^{\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}}}{1 + e^{\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}}} \quad (4.17)$$

ou, em forma matricial,

$$\pi = \frac{e^{X\beta}}{1 + e^{X\beta}} \quad (4.18)$$

Facilmente se verifica que a função $F : \mathbb{R} \rightarrow [0, 1]$, definida por

$$F(x) = \frac{\exp(x)}{1 + \exp(x)} \quad (4.19)$$

é uma função distribuição. Com efeito, trata-se da *função distribuição logística*, motivo pelo qual, o MLG definido pelo modelo binomial com função de ligação canónica *logit* é conhecido como modelo de regressão logística.

De referir que existem outras funções de ligação, tais como:

1. *Função probit* - Função de distribuição $N(0, 1)$;
2. *Função complementar log-log* - Função de distribuição Gumbel de mínimos;
3. *Função log-log* - Função de distribuição Gumbel de máximos.

No entanto, a transformação *logit* é usualmente preferida em relação às restantes, uma vez que:

1. Tem uma interpretação direta em termos do logaritmo do *odds*. Esta interpretação é, particularmente, útil na análise de dados epidemiológicos;
2. Os modelos baseados nesta transformação são apropriados para a análise de dados provenientes de um estudo retrospectivo;
3. Quando é usada esta transformação, é possível estimar diferenças na escala *logit* quer os dados sejam provenientes de um estudo retrospectivo quer de um estudo prospetivo. Esta propriedade não é, contudo, partilhada pelas restantes transformações.

O modelo de regressão logística é, portanto, dado por

$$\text{Logit}(\pi) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon \quad (4.20)$$

ou, em forma matricial,

$$\text{Logit}(\pi) = X\beta + \epsilon \quad (4.21)$$

Neste modelo, β_0 é o valor de $\ln\left(\frac{\pi}{1-\pi}\right)$ quando todos os $x_1 = x_2 = \cdots = x_k = 0$. $\beta_1, \beta_2 \cdots \beta_k$ são os coeficientes de regressão, isto é, a variação do $\text{logit}(\pi)$ quando $\Delta x_i = 1$. Note-se que, neste modelo, a variável resposta não é Y , ou $\mathcal{P}[Y = 1]$, mas sim o $\text{Logit}(\pi)$.

Dados agrupados ou em forma de proporções

Seja n_i o número de réplicas (indivíduos) para cada combinação $x_i^T = (x_{i1}, \dots, x_{ik})$ das variáveis explicativas. Então, a v.a. W_i , correspondente ao número de sucessos, é binomial de parâmetros n_i e π_i , isto é,

$$W_i \sim \text{Bin}(n_i, \pi_i) \quad (4.22)$$

O objetivo na regressão logística é o de investigar a relação entre a probabilidade da resposta, π_i , e as variáveis explicativas, $x_i^T = (x_{i1}, \dots, x_{ik})$. Para tal, não se utilizará W_i (frequências absolutas) como variável resposta, mas sim a proporção de sucessos $Y_i = \frac{W_i}{n_i}$ (frequências relativas). Deste modo, $E(Y_i) = \pi_i$ é, efetivamente, a probabilidade pretendida.

Tem-se que $n_i Y_i \sim \text{Bin}(n_i, \pi_i)$. Portanto, a f.m.p. da v.a. Y_i é dada por

$$\begin{aligned} f(y_i | \pi_i) = \mathcal{P}(Y_i = y_i) &= \mathcal{P}\left(\frac{\tilde{Y}_i}{n_i} = y_i\right) = \mathcal{P}(\tilde{Y}_i = n_i y_i) = \binom{n_i}{y_i n_i} \pi_i^{y_i n_i} (1 - \pi_i)^{n_i - y_i n_i} \\ &= \exp \left\{ y_i n_i \ln \pi_i + n_i (1 - y_i) \ln (1 - \pi_i) + \ln \binom{n_i}{y_i n_i} \right\} \\ &= \exp \left\{ y_i n_i (\ln \pi_i - \ln (1 - \pi_i)) + n_i \ln (1 - \pi_i) + \ln \binom{n_i}{y_i n_i} \right\} \\ &= \exp \left\{ n_i \left[y_i \ln \left(\frac{\pi_i}{1 - \pi_i} \right) + \ln (1 - \pi_i) \right] + \ln \binom{n_i}{y_i n_i} \right\} \end{aligned}$$

Facilmente se verifica que Y pertence à família exponencial, onde

$$\begin{aligned} \theta_i &= \ln \left(\frac{\pi_i}{1 - \pi_i} \right) \\ b(\theta_i) &= -\ln(1 - \pi_i) = \ln(1 + e^{\theta_i}) \quad ^3 \\ c(y_i, \phi) &= \ln \binom{n_i}{y_i n_i} \\ a(\phi) &= \frac{\phi}{\omega} = \frac{1}{n_i}, \phi = 1, \quad \omega = n_i \end{aligned}$$

De notar que

$$b'(\theta_i) = \frac{e^{\theta_i}}{1 + e^{\theta_i}} = \pi_i$$

$$\theta_i = \ln \left(\frac{\pi_i}{1 - \pi_i} \right) \Leftrightarrow e^{\theta_i} = \frac{\pi_i}{1 - \pi_i} \Leftrightarrow e^{\theta_i} (1 - \pi_i) = \pi_i \Leftrightarrow e^{\theta_i} = \pi_i + \pi_i e^{\theta_i} \Leftrightarrow e^{\theta_i} = \pi_i (1 + e^{\theta_i}) \Leftrightarrow \frac{e^{\theta_i}}{1 + e^{\theta_i}} = \pi_i.$$

$$\text{Portanto, } -\ln(1 - \pi_i) = -\ln \left(1 - \frac{e^{\theta_i}}{1 + e^{\theta_i}} \right) = -\ln \left(\frac{1 + e^{\theta_i} - e^{\theta_i}}{1 + e^{\theta_i}} \right) = -\ln(1 + e^{\theta_i})^{-1} = \ln(1 + e^{\theta_i})$$

$$b''(\theta_i) = \left(\frac{e^{\theta_i}}{1 + e^{\theta_i}} \right)' = \frac{e^{\theta_i}}{(1 + e^{\theta_i})^2} = \frac{\frac{\pi_i}{1 - \pi_i}}{\left(1 + \frac{\pi_i}{1 - \pi_i}\right)^2} = \pi_i(1 - \pi_i)$$

Portanto, de (4.8) e (4.9), tem-se que:

$$E(Y_i) = b'(\theta_i) = \pi_i \quad \text{e} \quad \text{var}(Y_i) = a(\phi)b''(\theta_i) = \frac{\pi_i(1 - \pi_i)}{n_i}$$

À semelhança do que acontecia com os dados binários (ou binomiais), $E(Y_i) = \pi_i$ e $\theta_i = \ln\left(\frac{\pi_i}{1 - \pi_i}\right)$, logo a ligação canónica é a função *logit*, isto é,

$$\theta_i = \eta_i = x_i^T \beta \Rightarrow \ln\left(\frac{\pi_i}{1 - \pi_i}\right) = x_i^T \beta$$

4.2.2 Coeficientes do Modelo

De forma a que seja possível interpretar corretamente as estimativas obtidas através de um modelo, é necessário, em primeiro lugar, conhecer o significado dos seus coeficientes. Para tal, ir-se-ão apresentar dois conceitos-chave na área da regressão logística.

Definição 4.2 (Odds).

Seja $\pi_i = \pi(x_i) = \mathcal{P}(Y = 1 | x_i^T)$, com $x_i^T = (x_{i1}, \dots, x_{ik})$.

Dá-se o nome de **odds** (também designado por chances ou rácio de verosimilhança) de um acontecimento, ao quociente entre a probabilidade de sucesso desse acontecimento, π_i , e a probabilidade de insucesso, isto é, as chances de se observar o sucesso ($Y=1$) relativamente ao insucesso ($Y=0$). De uma forma geral, o odds é dado por

$$\frac{\pi_i}{1 - \pi_i} \quad (4.23)$$

Por exemplo, se o odds de uma determinada doença é de 1 para 2 (1:2) isso significa que a ocorrência dessa doença está na razão de 1 para 2, ou seja, em cada 3 indivíduos, estima-se que 1 esteja doente.

Os coeficientes do modelo, $\beta_1, \beta_2, \dots, \beta_k$ não são fáceis de interpretar, pelo que é usual considerar a exponencial destes coeficientes, surgindo assim a definição 4.3.

Definição 4.3 (Odds ratio).

Quando dois conjuntos ($i = 1$ e $j = 2$) de dados binários são comparados dá-se o nome de **odds ratio** (ou rácio das chances) ao quociente

$$OR = \frac{\pi_1/(1 - \pi_1)}{\pi_2/(1 - \pi_2)} \quad (4.24)$$

Este quociente representa o rácio de chances do "sucesso" *versus* "insucesso" quando a variável independente X_i aumenta uma unidade relativamente às chances do "sucesso" *versus* "insucesso" quando X_i permanece constante. Este valor não é mais do que uma estimativa da razão de chances do "sucesso" *versus* "insucesso" por unidade da variável independente X_i . Ou seja, quando X_i varia uma unidade, as chances de obter "sucesso" variam β_i unidades.

1. Quando os *odds* nos 2 conjuntos de dados binários são iguais, o OR é igual a 1. Isto acontece quando as probabilidades de sucesso, dadas pelos *odds* em cada um dos conjuntos, são iguais;
2. Valores de OR menores que 1 sugerem que o *odds* no primeiro conjunto é menor do que no segundo. Por contraposição, valores de OR maiores do que 1 sugerem a situação contrária;
3. O *odds ratio* é uma medida da diferença entre duas probabilidades de sucesso que pode tomar qualquer valor positivo, ao contrário de $\pi_1 - \pi_2$ que varia no intervalo $(-1,1)$;
4. Se considerarmos o logaritmo tem-se

$$\ln OR = \ln \left(\frac{\pi_1}{1 - \pi_1} \right) - \ln \left(\frac{\pi_2}{1 - \pi_2} \right)$$

que não é mais do que a diferença dos *logit* nos dois conjuntos.

Significado dos coeficientes do modelo

No que se segue, considera-se $\pi = \pi(x) = P(Y = 1|x)$, com $x^T = (x_1, \dots, x_k)$ e somente a parte determinística é considerada.

Seja

$$\log \left(\frac{\pi}{1 - \pi} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k = \beta_j x_j + D, \quad (4.25)$$

onde $D = \beta_0 + \beta_1 x_1 + \dots + \beta_{j-1} x_{j-1} + \beta_{j+1} x_{j+1} + \dots + \beta_k x_k$.

A probabilidade π é dada por

$$\pi = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_k x_k)} = \frac{\exp \left(\beta_0 + \sum_{j=1}^k \beta_j x_j \right)}{1 + \exp \left(\beta_0 + \sum_{j=1}^k \beta_j x_j \right)} \quad (4.26)$$

isto é, $\pi = g^{-1}(x^T \beta)$.

Com base na equação (4.25) verifica-se que o parâmetro β_j corresponde à alteração produzida no *logit* pela alteração de uma unidade em x_j desde que:

1. x_j seja um fator com efeito linear;
2. a interação de x_j com os restantes fatores é nula;
3. os restantes fatores permaneçam constantes.

Em termos de *odds* vem:

$$\frac{\pi}{1 - \pi} = \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_j x_j + \cdots + \beta_k x_k)$$

Em termos de *odds ratio*, ao considerar-se uma alteração de d unidades em x_j tem-se

$$\frac{\text{odds}\{Y = 1|x_1, \dots, x_j + d, \dots, x_k\}}{\text{odds}\{Y = 1|x_1, \dots, x_j, \dots, x_k\}} = \frac{\exp(\beta_j(x_j + d)) \exp(D)}{\exp(\beta_j x_j) \exp(D)} = \exp(\beta_j d)$$

De forma a concretizar as noções atrás introduzidas, considere-se, sem perda de generalidade, o modelo de regressão logística com uma única covariável binária.

O modelo dado em (4.25) assume a forma

$$\text{logit}(Y = 1|x) = \log\left(\frac{\pi}{1 - \pi}\right) = \beta_0 + \beta_1 x \quad (4.27)$$

Tem-se que

$$\text{logit}(Y = 1|x = 0) = \ln\left(\frac{\pi(0)}{1 - \pi(0)}\right) = \beta_0 \quad (4.28)$$

$$\text{logit}(Y = 1|x = 1) = \ln\left(\frac{\pi(1)}{1 - \pi(1)}\right) = \beta_0 + \beta_1 \quad (4.29)$$

Donde se verifica que β_1 corresponde à diferença dos *logit* nos dois conjuntos, isto é,

$$\beta_1 = \ln OR = \ln\left(\frac{\pi(1)}{1 - \pi(1)}\right) - \ln\left(\frac{\pi(0)}{1 - \pi(0)}\right) \quad (4.30)$$

Desta forma, os *odds* são dados por

$$\text{odds}\{Y = 1|x = 0\} = \exp(\beta_0) \quad ; \quad \text{odds}\{Y = 1|x = 1\} = \exp(\beta_0 + \beta_1) \quad (4.31)$$

Consequentemente, o *odds ratio* vem dado por

$$OR = \frac{\text{odds}\{Y = 1|x = 1\}}{\text{odds}\{Y = 1|x = 0\}} = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \quad (4.32)$$

Interpretação

Seja y a variável em estudo, onde:

- morre corresponde a $y = 1$
- sobrevive (não morre) corresponde a $y = 0$

E x o fator em estudo, com:

- Tratamento B corresponde a $x = 1$
- Tratamento A corresponde a $x = 0$

Um *odds ratio* de 0.5, isto é,

$$OR = \frac{\pi(1)/(1 - \pi(1))}{\pi(0)/(1 - \pi(0))} = 0.5 \quad (4.33)$$

significa que o número de mortes no grupo que recebeu o tratamento B é metade (ou reduz 50%) do número de mortes no grupo que recebeu o tratamento A. Por outras palavras, pode-se ainda dizer que, a ocorrência de morte no grupo ao qual foi prescrito o tratamento A é duas vezes superior (ou aumenta 100%) à ocorrência de morte no grupo que recebeu o tratamento B.

4.2.3 Estimação dos Parâmetros

Nos modelos lineares generalizados, tal como nos restantes modelos, o objetivo primordial é a estimação dos parâmetros desconhecidos, neste caso, β e ϕ , e fazer inferência sobre os mesmos. O parâmetro de especial interesse e sobre o qual é fundamental fazer estimação é o β , uma vez que, o parâmetro ϕ é, geralmente, um parâmetro perturbador.

Uma vez que os erros do modelo logístico não seguem uma distribuição normal nem apresentam variância constante (a qual é função de $E(Y) = \pi$), ao contrário do que acontece com o modelo de regressão linear, não é possível obter estimativas ótimas dos coeficientes do modelo pelo método dos mínimos quadrados.

O método de estimação mais comum para a estimação de β é o Método da Máxima Verosimilhança (*Maximum Likelihood*). Este método estima os coeficientes de regressão que maximizam a probabilidade de encontrar realizações da variável dependente (y_1, y_2, \dots, y_n) amostradas, isto é, que maximizem a verosimilhança desses valores.

Para o modelo logístico, com dados agrupados (em forma de proporções), a função verosimilhança é dada por

$$\begin{aligned} L(\beta) &= \prod_{i=1}^n (y_i | \pi_i) = \prod_{i=1}^n \exp \left\{ n_i \left[y_i \ln \left(\frac{\pi_i}{1 - \pi_i} \right) + \ln(1 - \pi_i) \right] + \ln \left(\frac{n_i}{y_i n_i} \right) \right\} \\ &= \exp \left\{ \sum_{i=1}^n \left\{ n_i \left[y_i \ln \left(\frac{\pi_i}{1 - \pi_i} \right) + \ln(1 - \pi_i) \right] + \ln \left(\frac{n_i}{y_i n_i} \right) \right\} \right\} \end{aligned}$$

Uma vez que $\pi_i = \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}$, a verosimilhança pode escrever-se da seguinte forma:

$$L(\beta) = \exp \left\{ \sum_{i=1}^n \left\{ n_i \left[y_i \ln \left(\frac{\frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}}{1 - \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}} \right) + \ln \left(1 - \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)} \right) \right] + \ln \left(\frac{n_i}{y_i n_i} \right) \right\} \right\}$$

O máximo desta função ocorre para $\partial L / \partial \beta = 0$, com $\partial^2 L / \partial \beta^2 \leq 0$. No entanto, por conveniência de cálculo diferencial, já que é mais fácil, analiticamente, derivar uma soma do que um produto, ir-se-á considerar, de seguida, o logaritmo da verosimilhança, o qual se denomina, usualmente, por log-verosimilhança. Assim,

$$\ln L(\beta) = l(\beta) = \sum_{i=1}^n \left\{ n_i \left[y_i \ln \left(\frac{\frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}}{1 - \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}} \right) + \ln \left(1 - \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)} \right) \right] + \ln \left(\frac{n_i}{y_i n_i} \right) \right\}$$

Uma vez que

$$\begin{aligned} \ln(\pi_i) &= \ln \left(\frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)} \right) \\ &= \ln(\exp(x_i^T \beta)) - \ln(1 + \exp(x_i^T \beta)) \\ &= x_i^T \beta - \ln(1 + \exp(x_i^T \beta)) \end{aligned}$$

Tem-se, consequentemente,

$$\begin{aligned} \ln \left(\frac{\pi_i}{1 - \pi_i} \right) = x_i^T \beta &\Leftrightarrow \ln(\pi_i) - \ln(1 - \pi_i) = x_i^T \beta \\ &\Leftrightarrow -\ln(1 - \pi_i) = x_i^T \beta - \ln(\pi_i) \\ &\Leftrightarrow \ln(1 - \pi_i) = \ln(\pi_i) - x_i^T \beta \\ &\Leftrightarrow \ln(1 - \pi_i) = x_i^T \beta - \ln(1 + \exp(x_i^T \beta)) - x_i^T \beta \\ &\Leftrightarrow \ln(1 - \pi_i) = -\ln(1 + \exp(x_i^T \beta)) \end{aligned}$$

E, portanto, a log-verossimilhança simplifica-se, sendo dada por

$$\begin{aligned} l(\beta) &= \sum_{i=1}^n \left(n_i y_i \ln \left(\frac{\pi_i}{1 - \pi_i} \right) + n_i \ln (1 - \pi_i) + \ln \left(\frac{n_i}{y_i n_i} \right) \right) \\ &= \sum_{i=1}^n \left(n_i y_i x_i^T \beta - n_i \ln (1 + \exp (x_i^T \beta)) + \ln \left(\frac{n_i}{y_i n_i} \right) \right) \\ &= \sum_{i=1}^n l_i(\beta), \end{aligned}$$

onde

$$l_i(\beta) = n_i y_i x_i^T \beta - n_i \ln (1 + \exp (x_i^T \beta)) + \ln \left(\frac{n_i}{y_i n_i} \right)$$

O máximo de $l(\beta)$ ocorre com o vetor dos coeficientes β para o qual $\partial L(\beta)/\partial \beta = 0$. Porém, como $l(\beta)$ é uma transformação monótona de $L(\beta)$, o vetor que maximiza $l(\beta)$ é também o vetor que maximiza $L(\beta)$. Desta forma, os estimadores de máxima verossimilhança para β são a solução do sistema de equações abaixo indicado

$$\frac{\partial \ln L(\beta)}{\partial \beta_j} = 0 \Leftrightarrow \frac{\partial \sum_{i=1}^n l_i(\beta)}{\partial \beta_j} = 0, \quad j = 0, \dots, k.$$

Porém, este sistema de equações não possui solução analítica, pelo que β é estimado, iterativamente, por um algoritmo computacional⁴.

Apesar de não serem apresentados, é importante referir que os 2 métodos iterativos mais utilizados são o **Método de Newton-Raphson** e o **Método dos Mínimos Quadrados Ponderados (IRWLS)**. O método iterativo dos mínimos quadrados ponderados é semelhante ao método de Newton-Raphson, porém, utiliza a quantidade de informação de Fisher em vez da matriz Hessiana. Por este motivo, o método dos mínimos quadrados ponderados é também conhecido por **Scores de Fisher**.

4.2.4 Distribuições Assintóticas

Com o objetivo de inferir sobre o vetor de parâmetros $\hat{\beta}$ é necessário conhecer a distribuição de amostragem de $\hat{\beta}$. Todavia, nos MLG, em geral, não é possível obter as distribuições exatas, pelo que, ir-se-ão apresentar as distribuições assintóticas que se verificam quando os modelos satisfazem certas condições de regularidade.

⁴Na maioria dos softwares, para resolver este sistema de equações é usado o algoritmo de Newton-Raphson. No caso do software adotado, o R, é utilizado o Método Iterativo dos Mínimos Quadrados (*Fisher's Scoring Method*).

Os estimadores obtidos com o Método da Máxima Verosimilhança são eficientes, isto é, centrados e com variância mínima, e produzem estimativas com distribuição assintótica normal. Existe, contudo, um senão, uma vez que a inferência baseada na normalidade assintótica apenas é válida para amostras de grande dimensão. Apesar de não existir um critério rigoroso, uma amostra é considerada "grande" quando o número de observações é 10 a 20 vezes superior ao número de coeficientes do modelo.

Tem-se, assim, que:

1. O estimador de máxima verosimilhança de $\hat{\beta}$ é assintoticamente centrado, isto é, $E(\hat{\beta}) \simeq \beta$, e a matriz de variância-covariância é, aproximadamente, igual a $I^{-1}(\beta)$, isto é, $cov(\hat{\beta}) \simeq E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T] = I^{-1}(\beta)$;
2. Pelo Teorema de Limite Central (TLC), a distribuição assintótica de $\hat{\beta}$ é Normal, com valor médio β , e matriz de variância-covariância $I^{-1}(\beta)$ e escreve-se $\hat{\beta} \sim N(\beta, I^{-1}(\beta))$;
3. A estatística $(\hat{\beta} - \beta)^T I(\beta) (\hat{\beta} - \beta)$, conhecida como **estatística de Wald**, segue uma distribuição assintótica qui-quadrado com k graus de liberdade, isto é, segue um χ_k^2 ;
4. A distribuição assintótica de $\hat{\beta}_j$, $j = 1, \dots, k$, é Normal com valor médio β_j e variância $I_{jj}^{-1}(\beta)$ e escreve-se como $\hat{\beta} \sim N(\beta_j, I_{jj}^{-1}(\beta))$, onde $I_{jj}^{-1}(\beta)$ é o (j, j) -ésimo elemento de $I^{-1}(\beta)$.

Note-se que β é desconhecido e a matriz de informação de Fisher depende de β , pelo que, na prática, se substitui $I^{-1}(\beta)$ por $I^{-1}(\hat{\beta})$.

A **Matriz de informação de Fisher** é dada por

$$I(\beta) = E \left[-\frac{\partial^2 l}{\partial \beta \partial \beta^T} \right], \quad (4.34)$$

e é obtida através das segundas derivadas de l_i ,

$$= -E \left[\frac{\partial^2 l_i}{\partial \beta_j \partial \beta_l} \right] = E \left[\frac{\partial l_i}{\partial \beta_j} \frac{\partial l_i}{\partial \beta_l} \right] = \frac{x_{ij} x_{il}}{\text{var}(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2. \quad (4.35)$$

Repare-se que o (j, l) -ésimo elemento da matriz $I(\beta)$ é dado por:

$$= -\sum_{i=1}^n E \left[\frac{\partial^2 l_i}{\partial \beta_j \partial \beta_l} \right] = \sum_{i=1}^n \frac{x_{ij} x_{il}}{\text{var}(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2. \quad (4.36)$$

Em notação matricial, $I(\beta) = X^T \tilde{W} X$, onde

$$\widetilde{W} = \begin{pmatrix} \tilde{\omega}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\omega}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\omega}_n \end{pmatrix},$$

onde

$$\tilde{\omega} = \frac{1}{\text{var}(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 \quad (4.37)$$

4.2.5 Significância e Qualidade do Modelo de Regressão Logística

Uma vez ajustado o modelo de regressão logística, à semelhança do que acontece com a regressão linear, é necessário avaliar a significância e qualidade do ajustamento entre os dados observados e os dados ajustados, bem como a significância dos coeficientes de regressão logística. Para tal, recorrem-se a medidas que permitam aferir sobre a qualidade do modelo.

Com o objetivo de testar se as variáveis independentes são estatisticamente significativas para a explicação da variável dependente, é necessário realizar testes sobre o parâmetro β .

A maior parte dos testes de hipóteses que permitem testar a significância do modelo, são da forma

$$H_0 : C\beta = \xi \quad vs \quad H_1 : C\beta \neq \xi,$$

onde C é uma matriz $(q \times k)$, com $q \leq k$ e de característica completa k . ξ é vetor de dimensão q previamente especificado.

Em particular, estas hipóteses assumem a forma de:

1. Hipótese da nulidade de uma componente do vetor parâmetro, isto é, teste à nulidade de uma das covariáveis

$$H_0 : \beta_j = 0 \quad vs \quad H_1 : \beta_j \neq 0,$$

para algum $j \in 1, \dots, k$. Neste caso, $q = 1$, $C = (0, \dots, 0, 1, 0, \dots, 0)$, onde o 1 ocupa a j -ésima posição do vetor C e $\xi = 0$;

2. Hipótese da nulidade de r componentes de β . Suponhamos, por exemplo, que $H_1 : (\beta_1, \dots, \beta_k)^T = (0, \dots, 0)^T$. Então, $q = r$ e $C = (I_r O_{r \times (k-r)})$ e $\xi = 0_r$, onde I_r é a matriz identidade de ordem r , $O_{r \times (k-r)}$ é uma matriz de zeros de dimensão $(r \times (k-r))$ e 0_r é o vetor nulo de dimensão r .

Existem três testes distintos para testar as hipóteses anteriormente referidas. Contudo, apresentar-se-ão dois dos testes.

- **Teste de Wald:**

Para testar a hipótese (1) a estatística de teste de Wald é dada por:

$$W = (C\hat{\beta} - \xi)^T \left[CI^{-1}(\hat{\beta})C^T \right] (C\hat{\beta} - \xi), \quad (4.38)$$

cuja distribuição assintótica sob H_0 é um qui-quadrado com q graus de liberdade, isto é, χ_q^2 .

A região de rejeição é dada por $\{W_{obs} > \chi_q^{1-\alpha}\}$.

Este teste tem grande utilidade quando se pretende testar se determinada covariável é significativa para o modelo e, neste caso, o teste é dado por

$$H_0 : \beta_j = 0 \quad vs \quad H_1 : \beta_j \neq 0,$$

pelo que a estatística de teste vem dada por

$$W = \frac{\hat{\beta}_j}{\sigma_{jj}} \quad (4.39)$$

W segue um $\chi_{(1)}^2$ e σ_{jj} corresponde ao j -ésimo elemento da diagonal da matriz $I^{-1}(\hat{\beta})$.

A região de rejeição é dada por $\{W_{obs} > \chi_1^{1-\alpha}\}$.

- **Teste de Razão de Verossimilhanças:**

Considere-se $\tilde{\beta}$ o estimador de máxima verossimilhança restrito, isto é, $\tilde{\beta}$ é o valor de β que maximiza a verossimilhança sujeita à hipótese $H_0 : C\beta = \xi$.

A estatística de teste à significância do modelo compara a verossimilhança do modelo só com a constante, usualmente designado por modelo nulo ou modelo reduzido (isto é, nenhuma das covariáveis tem poder preditor e, portanto, $\text{logit}(\pi) = \beta_0$) com a verossimilhança do modelo com todas as variáveis explicativas, denominado de modelo completo.

A estatística de razão de verossimilhanças ou de Wilks é dada por

$$\Lambda = -2 \ln \frac{\max_{H_0} L(\beta)}{\max_{H_0 \cup H_1} L(\beta)} = -2 \{ \ln L(\tilde{\beta}) - \ln L(\hat{\beta}) \}, \quad (4.40)$$

onde $\tilde{\beta}$ e $\hat{\beta}$ são os estimadores de máxima verossimilhança de β , sob H_0 e sob $H_0 \cup H_1$, respetivamente. A distribuição assintótica de λ é, sob certas condições

de regularidade e sob H_0 , um qui-quadrado cujo número de graus de liberdade (g.l.) é dado pela diferença entre o número de parâmetros a estimar sob $H_0 \cup H_1$ e o número de parâmetros a estimar sob H_0 .

A região de rejeição é, portanto, dada por $\{\Lambda_{obs} > \chi_1^{1-\alpha}\}$.

Esta estatística de teste é um indicador de mediocridade do ajustamento do modelo aos dados, isto é, quanto maior for o valor observado da estatística de teste, pior é o ajustamento. Por outro lado, se o valor observado da estatística de teste, Λ_{obs} , for igual a 0, o ajustamento é perfeito.

A utilidade do teste de Wilks traduz-se na possibilidade de comparação entre modelos encaixados.

Importante referir que este teste apenas permite concluir que o modelo é significativo, porém, não nos indica que todas as variáveis incluídos no modelo são significativas. Assim sendo, este teste por si só não permite aferir que o ajustamento é bom.

Desta forma, para testar a significância do ajustamento do modelo completo é necessário formular as seguintes hipóteses

H_0 : O modelo ajusta-se bem aos dados *vs* H_1 : O modelo não se ajusta aos dados

• **Estatística de Pearson Generalizada - X^2** :

A estatística de teste clássica é o Qui-quadrado de Pearson (Hosmer and Lemeshow, 2000), a qual é definida, de um modo geral, por

$$X^2 = \sum_{i=1}^n \frac{(\omega_i(y_i - \hat{\mu}_i))^2}{\hat{\phi}V(\hat{\mu}_i)} \quad (4.41)$$

Para o modelo logístico, com $Y \sim \text{Bin}(n_i, \pi_i)$, a estatística é dada por

$$\begin{aligned} X^2 &= \sum_{i=1}^n \frac{(y_i - n_i\pi_i)^2}{n_i\pi_i(1 - \pi_i)} \\ &= \sum_{i=1}^n \frac{(y_i - n_i\pi_i)^2}{n_i\pi_i} + \sum_{i=1}^n \frac{((n_i - y_i) - n_i(1 - \pi_i))^2}{n_i(1 - \pi_i)} \\ &= \sum_{i=1}^n \frac{(o_i - e_i)^2}{e_i}, \end{aligned} \quad (4.42)$$

onde o_i e e_i são, respetivamente, o número de sucessos observados e de sucessos esperados na célula i .

Esta estatística, sob H_0 e para amostras grandes, apresenta uma distribuição assintótica de um χ^2_{n-k-1} , onde k é o número de parâmetros estimados.

Portanto, para o modelo logístico, a estatística de Pearson generalizada coincide com a estatística original de Pearson.

A título de curiosidade, refira-se que acontece a mesma situação quando considerado o modelo Poisson.

• Função Desvio - Deviance $D(y; \hat{\mu})$:

Alguns autores, como por exmplo, McCullagh e Nelder (1989) utilizam a estatística de teste *Deviance* ou Desvio.

Esta medida é baseada na estatística de razão de verosimilhanças e avalia a discrepância entre o modelo saturado (S), isto é, o modelo com tantos parâmetros quanto observações, e o modelo corrente (C).

A estatística *Deviance* é dada por

$$D^*(y; \mu) = -2\{\ln L_C(\tilde{\beta}) - \ln L_S(\hat{\beta})\}, \quad (4.43)$$

onde $\ln L_S(\hat{\beta})$ corresponde à verosimilhança máxima alcançada para um ajustamento exato, no qual existem tantos parâmetros quanto observações, normalmente. Por outro lado, $\ln L_C(\tilde{\beta})$ representa a verosimilhança do modelo corrente.

Considerando o modelo logístico com dados em forma de proporções, como já fora mencionado, para o modelo saturado (S), cada parâmetro π_i é estimado com base no valor observado, isto é, $\hat{\pi}_i = y_i$. Por outro lado, para o modelo corrente (M), existem q parâmetros, com $q < n$ e os valores de $\hat{\pi}_i$ são estimados com recurso aos

valores ajustados, isto é, $\hat{\pi}_i = \hat{y}_i$. Desta forma, a função desvio vem dada por:

$$\begin{aligned}
 D(y; \mu) &= 2l(y; y) - 2l(\mu; y) = -2\{l(\mu; y) + l(y; y)\} \\
 &= -2 \sum_{i=1}^n \left\{ \left\{ n_i \left[y_i \ln \left(\frac{\hat{y}_i}{1 - \hat{y}_i} \right) + \ln(1 - \hat{y}_i) \right] + \ln \left(\frac{n_i}{y_i n_i} \right) \right\} - \right. \\
 &\quad \left. - \left\{ n_i \left[y_i \ln \left(\frac{y_i}{1 - y_i} \right) + \ln(1 - y_i) \right] + \ln \left(\frac{n_i}{y_i n_i} \right) \right\} \right\} \\
 &= -2 \sum_{i=1}^n \left\{ \left\{ n_i \left[y_i \ln \left(\frac{\hat{y}_i}{1 - \hat{y}_i} \right) + \ln(1 - \hat{y}_i) \right] \right\} - \right. \\
 &\quad \left. - \left\{ n_i \left[y_i \ln \left(\frac{y_i}{1 - y_i} \right) + \ln(1 - y_i) \right] \right\} \right\} \\
 &= -2 \sum_{i=1}^n \{ \{ n_i y_i \ln(\hat{y}_i) - n_i y_i \ln(1 - \hat{y}_i) + n_i \ln(1 - \hat{y}_i) \} - \\
 &\quad - \{ n_i y_i \ln(y_i) - n_i y_i \ln(1 - y_i) + n_i \ln(1 - y_i) \} \} \\
 &= -2 \sum_{i=1}^n \left\{ n_i y_i \ln \left(\frac{\hat{y}_i}{y_i} \right) - n_i y_i \ln \left(\frac{1 - \hat{y}_i}{1 - y_i} \right) + n_i \ln \left(\frac{1 - \hat{y}_i}{1 - y_i} \right) \right\} \\
 &= -2 \sum_{i=1}^n \left\{ n_i y_i \ln \left(\frac{\hat{y}_i}{y_i} \right) + (n_i - n_i y_i) \ln \left(\frac{1 - \hat{y}_i}{1 - y_i} \right) \right\} \\
 &= -2 n_i \sum_{i=1}^n \left\{ y_i \ln \left(\frac{\hat{y}_i}{y_i} \right) + (1 - \hat{y}_i) \ln \left(\frac{1 - \hat{y}_i}{1 - y_i} \right) \right\}
 \end{aligned}$$

De acordo com o desvio obtido, é possível aferir sobre a qualidade do ajustamento, em que:

- **Deviance pequena:** Significa que a explicação do modelo ajustado é praticamente igual à do modelo completo, pelo que se pode concluir que se trata de um bom ajustamento. Neste caso, deve-se optar pelo modelo ajustado, uma vez que tem menos variáveis independentes, tornando-se num modelo parcimonioso;
- **Deviance grande:** Indica que a explicação do modelo ajustado é pobre, invalidando a utilização do mesmo.

A noção de desvio é frequentemente utilizada quando se pretende comparar modelos encaixados ou aninhados.

Sejam M_0 e M dois modelos onde M_0 é um submodelo de M com $q < k$, isto é, $M_0 \subset M$. Se ϕ for conhecido, a estatística de teste de razão de verosimilhanças para comparar estes dois modelos é dada por

$$\frac{D_{M_0} - D_M}{\phi} \tag{4.44}$$

A esta estatística de teste dá-se o nome de **diferença de deviances**, onde D_{M_0} é o desvio para o modelo M_0 e D_M é o desvio para o modelo M . Sob a validade do modelo M_0 , a distribuição assintótica da estatística de teste apresentada em (4.44) é um χ^2_{k-q} .

A função desvio (ou *deviance*) é mais utilizada do que a estatística de Pearson Generalizada, uma vez que o valor do desvio diminui quando se adiciona variáveis ao modelo e tal não acontece com a estatística de Pearson.

Todavia, se o modelo apresentar variáveis quantitativas, caso em que o número de células é aproximadamente igual ao número de elementos da amostra, as duas estatísticas atrás apresentadas (Estatística de Pearson Generalizada e Desvio) não podem ser usadas com dados individuais. Perante este cenário, a distribuição χ^2 da estatística de teste não é, geralmente, válida (relembre-se que existem condições para a aplicação do teste do Qui-quadrado, em particular, a condição que exige que 80% das células com valor esperado maior ou igual a 5). Importante referir que se trata de uma regra empírica e a qual não inviabiliza a utilização desta estatística.

- **Estatística de Hosmer-Lemeshow** :

Hosmer and Lemeshow (2000) propuseram uma outra estatística que permite testar o ajustamento do modelo aos dados, baseada na estimativa das probabilidades de sucesso de cada uma das n observações. A estatística de teste é obtida com um teste de Qui-quadrado a uma tabela de contingência $2 \times g$. Esta tabela é construída classificando as duas classes (ou categorias) da variável dependente dicotômica por g grupos dicotômicos definidos pelos decis das probabilidades de "sucesso" (isto é, $[0; 0, 1[, [0, 1; 0, 2[, \dots, [0, 9; 1]$) estimadas pelo modelo para cada uma das observações. Geralmente, $g = 10$, porém, se o modelo tiver variáveis independentes qualitativas, g pode ser inferior ou igual a 10, em função do número de células distintas resultantes do cruzamento das classes das variáveis independentes. O número de "sucessos" da variável dependente em cada classe definida pelos decis (o_i) e os valores esperados nessas classes, usando a média das probabilidades estimadas para cada um dos grupos da variável dependente e cada uma das classes decílicas ($e_i = n_i \pi_i$) são usadas no cálculo da estatística de teste:

$$X_{HL}^2 = \sum_{i=1}^g \frac{(o_i - e_i)^2}{e_i} \quad (4.45)$$

Para amostras de grande dimensão, a estatística de teste, anteriormente apresentada, tem distribuição assintótica $\chi^2_{(g-2)}$.

Análise de resíduos

Para avaliar a qualidade do modelo é importante analisar os resíduos. Os resíduos, assim como características construídas a partir deles, constituem a base de ferramentas de diagnóstico utilizadas para verificar a validade dos pressupostos do modelo.

Existem várias ferramentas para realizar a análise de resíduos, das quais se destacam as ferramenteas gráficas, tendo em conta que são intuitivas.

Os tipos de resíduos mais utilizados no MLG são:

- Resíduos correntes:

$$r_i = y_i - \hat{\mu}_i, \quad i = 1, \dots, n, \quad (4.46)$$

onde y_i são os valores observados e μ_i os valores ajustados. Uma vez que a variável resposta é binária (assume os valores 0 e 1) e os valores ajustados variam no intervalo $[0,1]$, os resíduos do modelo variam no intervalo $[-1,1]$. Quando $r_i > 0$, $y_i = 1$ e, para $r_i < 0$, $y_i = 0$. Para $r_i = 0$, o ajustamento é perfeito, isto é, $y_i = \hat{\mu}_i$ (os valores observados são iguais aos valores ajustados).

- Resíduos de Pearson:

$$r_{pi} = \frac{(y_i - \hat{\mu}_i)\omega_i}{\sqrt{\hat{\phi}(\text{var}(\hat{\mu}_i))}}, \quad i = 1, \dots, n. \quad (4.47)$$

Para o modelo logístico, com $Y_i \sim \text{Bin}(n_i, \pi_i)$, tem-se, como já fora mencionado,

$$\hat{\mu}_i = n_i \pi_i, \quad V(\mu_i) = n_i \pi_i (1 - \pi_i) \quad \text{e} \quad \phi = \omega = 1$$

E, portanto, os resíduos de Pearson, para o modelo logístico, são dados por

$$r_{pi} = \frac{(y_i - \hat{\mu}_i)\omega_i}{\sqrt{\hat{\phi}(\text{var}(\hat{\mu}_i))}} = \frac{(y_i - \hat{\mu}_i)}{\sqrt{\text{var}(\hat{\mu}_i)}} = \frac{(y_i - n_i \pi_i)}{\sqrt{n_i \pi_i (1 - \pi_i)}}, \quad i = 1, \dots, n. \quad ^5 \quad (4.48)$$

- Resíduos de Pearson padronizados:

$$r_{pi}^P = \frac{(y_i - \hat{\mu}_i)\omega_i}{\sqrt{\text{var}(\hat{\mu}_i)(1 - h_{ii})}}, \quad i = 1, \dots, n, \quad (4.49)$$

⁵Para o modelo logístico com dados agrupados, isto é, na forma de proporções o procedimento é semelhante, considerando $\hat{\mu}_i = \pi_i$, $V(\hat{\mu}_i) = \frac{\pi_i(1 - \pi_i)}{n_i}$, $\phi = 1$ e $\omega = n_i$.

onde h_{ii} é o i -ésimo elemento da diagonal da matriz de projeção generalizada \mathbf{H} (também chamada de *hat matrix*, uma vez que $\hat{\mu}_i = \mathbf{H}y$). Este termo pode ser calculado através da expressão $h_{ii} = x_i^T (X^T X)^{-1} x_i$. A desvantagem dos resíduos de Pearson é que a sua distribuição é, geralmente, bastante assimétrica para modelos não Normais, do qual é exemplo o modelo de regressão logística.

Tal como referido anteriormente, a análise de resíduos do modelo é feita, principalmente, com recurso a ferramentas gráficas, das quais se destacam as seguintes:

- **Scatterplot dos resíduos** - Conhecido como diagrama de dispersão. Este gráfico permite verificar se os resíduos não apresentam qualquer tipo de padrão e se os mesmos se encontram bem distribuídos no intervalo $[-2,2]$. De acordo com a literatura existente, 95% dos resíduos devem encontrar-se nesse intervalo;
- **Resíduos vs Valores ajustados** - Permite avaliar se a variância dos resíduos é não constante ou, por outras palavras, permite verificar a existência de heterocedasticidade. É assumido que os resíduos são independentes dos valores ajustados, querendo isto dizer que a correlação entre resíduos e valores preditos deve tomar o valor 0;
- **Q-Q plot Normal dos resíduos** - Permite avaliar se os resíduos seguem uma distribuição Normal(0,1), através da comparação entre os quantis teóricos e amostrais;
- **Histograma dos resíduos** - Permite avaliar a simetria (ou falta dela) dos resíduos, ajudando, eventualmente, na deteção de padrões nos resíduos.

4.2.6 Pseudo- R^2

Uma outra forma de avaliar a qualidade do modelo é, à semelhança da regressão linear, a de avaliar a dimensão do efeito do modelo. Na regressão linear, a dimensão do efeito linear da(s) variável(eis) independente(s) sob a variável resposta é avaliada com recurso à medida R^2 (coeficiente de determinação). Contudo, na regressão logística, não é possível calcular o R^2 já que a variância da variável resposta depende da probabilidade de ocorrência dos seus valores. Assim, os R^2 para os diferentes modelos não podem ser comparados diretamente, a menos que as funções densidade das variáveis dependentes nos diferentes modelos sejam iguais.

Surgiu, assim, a necessidade de encontrar uma medida relacionada com o coeficiente de determinação para os modelos com resposta binária e multinominal. Várias medidas

foram propostas nas últimas três décadas das quais se destacam as sugeridas por McFadden (1973), McKelvey and Zavoina (1975), Maddala (1983), Agresti (1986), Cox and Snell (1989), Nagelkerke (1991), Cox and Wermuch (1992), Ash and Shwartz (1999) e Zheng and Agresti (1999).

Ir-se-ão apresentar três dos pseudo- R^2 sugeridos até à data, onde $\ln L_N$ corresponde à log-verosimilhança do modelo nulo, isto é, sem variáveis preditoras e $\ln L_S$ representa a log-verosimilhança do modelo saturado ou completo, isto é, o modelo no qual são consideradas todas as covariáveis.

Pseudo- R^2 de McFadden

McFadden (1973) sugeriu uma medida que compara o modelo sem qualquer variável preditora, denominado de modelo nulo e o modelo que considera todos os preditores, dada por

$$R_{MF}^2 = 1 - \frac{\ln L_S}{\ln L_N} \quad (4.50)$$

Pseudo- R^2 de Cox & Snell

Cox and Snell (1989) sugeriram outra expressão para calcular o pseudo- R^2 , a qual é dada por

$$R_{CS}^2 = 1 - e^{-\frac{2(\ln L_S - \ln L_N)}{n}} \quad (4.51)$$

Esta medida pode ainda ser reescrita em função da estatística de razão de verosimilhanças ($\Lambda = -2\{\ln L_N - \ln L_S\}$), sendo dada por $R_{CS}^2 = 1 - e^{-\Lambda/n}$.

Note-se que esta estatística nunca atinge o valor 1, mesmo quando o ajustamento é perfeito.

Pseudo- R^2 de Nagelkerke

Baseado no pseudo- R^2 sugerido por Cox and Snell (1989), Nagelkerke (1991) propôs uma correção ao R_{CS}^2 de modo a que este varie no intervalo $[0,1]$:

$$R_N^2 = \frac{R_{CS}^2}{1 - e^{-\frac{2 \ln L_N}{n}}} \quad (4.52)$$

O R_N^2 já pode tomar o valor 1, contudo, ao contrário do que acontece na regressão linear, os valores intermédios não são interpretados como sendo a percentagem da variabilidade da variável dependente que é explicada pelo modelo. Hosmer and Lemeshow (2000) referem que os pseudo- R^2 na regressão logística são baseados na comparação do modelo ajustado com o modelo nulo e, consequentemente, não são medidas da variabilidade explicada pelo modelo.

Shatland *et al.* (2002) estudaram a utilização de vários pseudo- R^2 na regressão logística e concluíram que a estatística que apresenta uma melhor interpretação é o pseudo- R^2 de McFadden. De acordo com estes autores, o R_{MF}^2 pode ser interpretado com a proporção da redução da log-verosimilhança do modelo nulo, relativamente ao modelo completo. Por outras palavras, esta medida corresponde ao rácio entre o ganho de informação estimada pelo modelo completo em comparação com o modelo nulo e o ganho da informação, potencialmente, recuperável por um modelo saturado.

4.2.7 Capacidade Preditiva do Modelo

4.2.7.1 Sensibilidade e Especificidade

Uma vez terminada a análise de resíduos é fundamental avaliar a capacidade discriminatória do modelo.

Como os valores preditos dados pelo modelo são contínuos (pertencem ao intervalo $[0,1]$), é necessário definir um ponto de corte (*cut-off*), que permite classificar e contabilizar o número de predições em que a variável resposta toma os valores 0 e 1. O ponto de corte é tal que, se a propensão dada pelo modelo para uma dada observação for inferior a este valor, assume-se que a observação se encontra na categoria 0 da variável resposta. Por outro lado, se a propensão for superior ao ponto de corte considerado, assume-se que a observação pertence à categoria 1 da variável resposta. Uma vez fixado o ponto de corte é possível dividir as propensões em duas categorias: abaixo do *cut-off* (0) e acima do *cut-off* (1).

Tendo em conta que, para o conjunto de observações consideradas, é conhecido o valor real da resposta (ou valor observado), é possível compará-lo com a categoria da predição feita pelo modelo. Desta forma, existem 4 casos possíveis:

- A categoria de predição é 1 e o valor observado da variável resposta é também 1. A este grupo dá-se o nome de verdadeiros positivos (VP);
- A categoria de predição é 1 e o valor observado da variável resposta é 0. A esta situação dá-se o nome de falsos positivos (FP);

- A categoria de predição é 0 e o valor observado da variável resposta é 1. A este grupo dá-se o nome de falsos negativos (FN);
- A categoria de predição é 0 e o valor observado da variável resposta é 0. A este grupo dá-se o nome de verdadeiros negativos (VN);

Esta comparação pode ser sintetizada na tabela apresentada na figura 4.3, a qual é, usualmente, designada por *Confusion Matrix*, não possuindo uma tradução consensual para a língua portuguesa.

		Valores observados	
		positivo (1)	negativo (0)
Valores preditos	positivo (1)	VP (verdadeiro positivo)	FP (falso positivo)
	negativo (0)	FN (falso negativo)	VN (Verdadeiro negativo)

Figura 4.3: *Confusion Matrix* para um modelo de regressão logística

Com base nos conceitos apresentados na tabela apresentada na figura 4.3, é possível determinar as seguintes medidas:

- **Precisão** (do inglês *accuracy*)- Representa a proporção de predições corretas. Esta medida é suscetível a conjuntos de dados desequilibrados, uma vez que não tem em conta o número de elementos que pertencem a cada categoria. Desta forma pode facilmente conduzir a conclusões incorretas sobre o desempenho do modelo.

É dada por:

$$ACC = \frac{N^{\circ} \text{ total de acertos}}{N^{\circ} \text{ de observações em estudo}} = \frac{VP + VN}{N}, \quad (4.53)$$

onde N representa o número total de observações.

- **Sensibilidade** - Taxa de verdadeiros positivos, isto é, corresponde à percentagem de classificações corretas na classe de referência "1-Sucesso" da variável resposta (o sujeito tem a característica que se pretende modelar e o modelo prevê essa característica corretamente). É dada por:

$$SENS = \frac{N^{\circ} \text{ de acertos positivos}}{N^{\circ} \text{ de observações positivas}} = \frac{VP}{VP + FN} \quad (4.54)$$

- **Especificidade** - Taxa de verdadeiros negativos, ou seja, corresponde à percentagem de classificações corretas na classe de "0-insucesso" do modelo (o indivíduo não possui a característica a modelar e o modelo prevê corretamente que o sujeito não possui essa mesma característica). É dada por:

$$ESP = \frac{N^{\circ} \text{ de acertos negativos}}{N^{\circ} \text{ de observações negativos}} = \frac{VN}{VN + FP} \quad (4.55)$$

- **Eficiência** - Corresponde à média aritmética (ou semi-soma) da sensibilidade e da especificidade e como tal é dada por:

$$EFF = \frac{SENS + ESP}{2} \quad (4.56)$$

Repare-se que, na prática, a sensibilidade variam em sentidos opostos (adiante ilustrar-se-á o porquê), uma vez que, geralmente, quando um modelo é muito sensível a positivos (categoria 1 da variável resposta), tendo a gerar muitos falsos positivos e vice-versa.

De notar que a sensibilidade e a especificidade não são calculadas sobre os mesmos indivíduos, uma vez que no cálculo da sensibilidade apenas são utilizados os indivíduos que se encontram na categoria 1 da variável de interesse e na especificidade utilizam-se somente os indivíduos que se encontram na categoria 0. Sendo assim, estas medidas são independentes.

Por analogia, na biomedicina, onde estas medidas são frequentemente utilizadas em diagnósticos clínicos, pode-se definir sensibilidade como sendo a probabilidade condicionada de um diagnóstico positivo sabendo que o indivíduo possui a doença e a especificidade como sendo a probabilidade condicionada de um diagnóstico negativo sabendo que o indivíduo não tem a doença.

De acordo com Marôco (2010), pode-se classificar a capacidade do modelo preditivo de acordo com a sensibilidade e especificidade apresentadas na tabela 4.1.

Tabela 4.1: Capacidade preditiva do modelo de acordo com a sensibilidade e especificidade

Percentagem	Capacidade preditiva
>0,8	Boas Capacidades preditivas
]0,5;0,8[Capacidades preditivas razoáveis
<0,5	Capacidades prediitivas medíocres

4.2.7.1 Curva ROC

Aliada aos conceitos de especificidade e sensibilidade, encontra-se a Curva ROC (*Receiver Operating Characteristic*), que é uma poderosa ferramenta para medir e especificar problemas no desempenho do diagnóstico em medicina, uma vez que permite estudar a variação da sensibilidade e especificidade para diferentes valores de corte. O desempenho de um teste é, usualmente, demonstrado pela curva ROC.

As curvas ROC foram desenvolvidas no campo das comunicações com o objetivo de estudar a relação sinal-ruído. O sinal era interpretado como verdadeiro positivo (sensibilidade) e o ruído como falso positivo (1-especificidade). Mais concretamente, foram desenvolvidas na deteção de sinais eletrónicos e problemas com radares, durante a Segunda Guerra Mundial (Zweig and Campbell, 1993).

Na década de 60, as curvas ROC foram utilizadas em psicologia experimental e, nos anos 70, a metodologia rapidamente se disseminou em vários ramos da pesquisa biomédica, área na qual a sua aplicação tem como objetivo auxiliar a classificação de indivíduos em doentes ou não doentes.

A curva ROC traduz-se na representação gráfica da sensibilidade, no eixo das ordenadas, e, no eixo das abcissas, do complementar da especificidade, isto é, $(1 - ESP)$.

Como se pode ver pela figura 4.4, um classificador perfeito corresponderia a uma linha horizontal sobre o gráfico, porém dificilmente será alcançada. Na prática, curvas consideradas boas encontram-se entre a linha diagonal (classificador aleatório) e a linha perfeita, onde, quanto maior a distância à linha diagonal, melhor o desempenho do modelo. Tal como referido, a linha diagonal corresponde a um classificador aleatório, isto é, no qual a probabilidade de ser classificado positivo é igual à probabilidade de ser classificado como negativo (0,5).

Grande parte dos autores, utiliza a área abaixo da curva (*AUC*, do inglês *Area Under Curve*) como um dos índices mais utilizados para avaliar a qualidade da curva. A área sob a curva varia entre 0 e 1. Hosmer and Lemeshow (2000) apresentam valores indicativos da área sob a curva ROC que podem ser utilizados para classificar o poder discriminante do modelo de regressão e que, numa regra de aplicação geral, são dados pela tabela

De notar que a área sobre a curva pode ser obtida através de diversos métodos de integração numérica, como por exemplo, a método do trapézio.

Por fim, a partir da curva ROC, é possível determinar o ponto de corte que maximiza o desempenho do modelo, denominado, usualmente, por *cut-off* ótimo. Este *cut-off* traduz o compromisso existente entre o número de verdadeiros positivos e o número de falsos

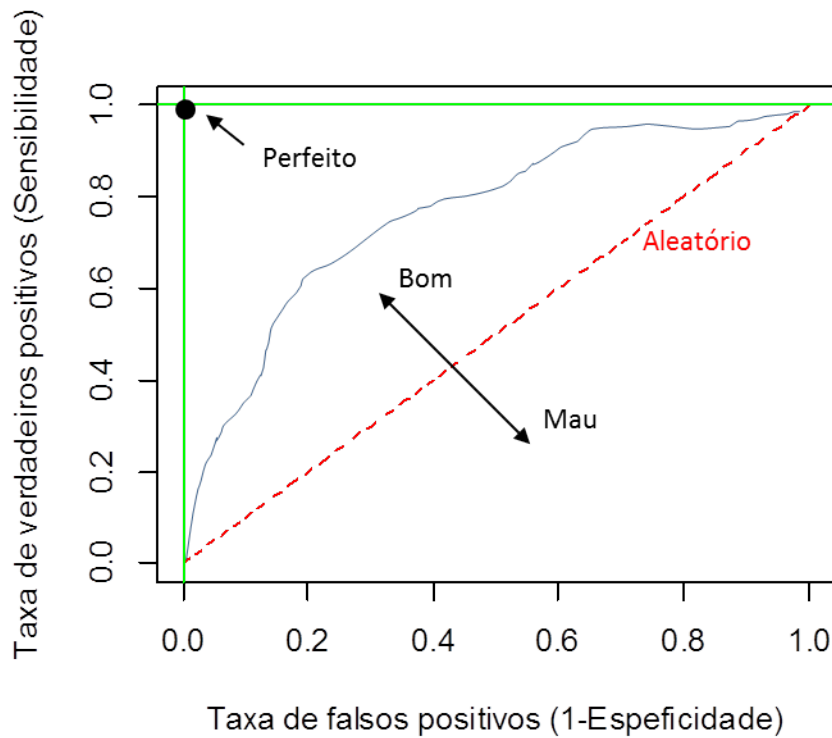


Figura 4.4: Curva ROC.

Tabela 4.2: Valores indicativos da área sob a curva ROC

Área sob a curva ROC	Poder discriminante do modelo
0,5	Sem poder discriminativo
]0,5;0,7[Discriminação fraca
]0,7;0,8[Discriminação razoável
]0,8; 0,9[Discriminação boa
$\geq 0,9$	Discriminação ótima

positivos.

Dos métodos existentes para determinar o *cut-off* ótimo, no presente trabalho utilizar-se-á o método denominado de *Youden's index* (usualmente representado por J). Este índice corresponde à diferença entre a taxa de verdadeiros positivos e a taxa de falsos positivos. A maximização deste índice corresponde à determinação do *cut-off* ótimo. De acordo com a sua definição, J é a distância vertical entre a curva ROC e a linha diagonal referente à aleatoriedade (a vermelho na figura 4.4).

Se $R(x)$ for a função que descreve a curva ROC, com $x = 1 - ESP$, $J(x) = R(x) -$

x . Quando J é máximo, $J'(x) = 0$, onde J' corresponde à derivada de 1ª ordem de J . Consequentemente, $R'(x) = 1$, com R' a 1ª derivada de R . Uma vez que $R'(x) = 1$, a tangente à curva ROC é paralela à linha diagonal, que possui declive 1. Desta forma, determinar o ponto de corte ótimo corresponde simplesmente à deslocação da linha diagonal até que a mesma se torne tangente à curva. Ao encontrar esta reta, obtém-se o par (sensibilidade, 1-especificidade) correspondente ao *cut-off* ótimo.

4.2.8 Intervalos de Confiança

O intervalo de confiança (IC) para os parâmetros β_j , $j = 1, \dots, k$, ao nível de confiança, de, aproximadamente, $(1 - \alpha) \times 100\%$ podem ser facilmente obtidos através da estatística de Wald e são dados por

$$\left(\hat{\beta}_j \pm z_{1-\alpha/2} \sqrt{I^{-1}(\hat{\beta})} \right), \quad (4.57)$$

onde $z_{1-\alpha/2}$ é o quantil de $1 - \alpha/2$ da distribuição Normal padrão.

Para o vetor β de dimensão k ,

$$\left(\hat{\beta} - \beta \right)^T I(\hat{\beta}) \left(\hat{\beta} - \beta \right) \leq \chi_{1-\alpha, k}^2, \quad (4.58)$$

onde $\chi_{1-\alpha, k}^2$ é o quantil $1 - \alpha$ do χ^2 com k g.l., é dado o elipsóide de confiança para β .

O intervalo de confiança ao nível de, aproximadamente, $(1 - \alpha) \times 100\%$ para $d\beta_j$ é dado por

$$\left(d\hat{\beta}_j \pm z_{1-\alpha/2} d\sqrt{\text{var}(\hat{\beta}_j)} \right), \quad (4.59)$$

onde $z_{1-\alpha/2}$ é o quantil $(1 - \alpha/2)$ da normal padrão.

Consequentemente, os intervalos de confiança ao nível de, aproximadamente, $(1 - \alpha) \times 100\%$ para $\exp(\beta_j)$ (*odds ratio*) e $\exp(d\beta_j)$ são dados, respetivamente, por

$$\left(\exp \left(\hat{\beta}_j \pm z_{1-\alpha/2} \sqrt{\text{var}(\hat{\beta}_j)} \right) \right) \quad (4.60)$$

$$\left(\exp \left(d\hat{\beta}_j \pm z_{1-\alpha/2} d\sqrt{\text{var}(\hat{\beta}_j)} \right) \right), \quad (4.61)$$

O intervalo de confiança ao nível de, aproximadamente, $(1 - \alpha) \times 100\%$ para o *logit* é dado por

$$\left(g(\hat{\pi}_i \pm z_{1-\alpha/2} \sqrt{\text{var}(g(\hat{\pi}_i))} \right), \quad (4.62)$$

onde

$$\text{var}(g(\hat{\pi}_i)) = \sum_{j=1}^k x_{ij}^2 \text{var}(\hat{\beta}_j) + \sum_{j=1}^k \sum_{l=1, j \neq l}^k x_{ij} \text{cov}(\hat{\beta}_j, \hat{\beta}_l) = X_i^T I^{-1}(\hat{\beta}) X_i \quad (4.63)$$

Capítulo 5

Análise de Correspondências

5.1 Introdução

A Análise de Correspondências (AC), tal como é conhecida hoje em dia, percorreu um longo caminho em 30 anos, desde que foi denominada de "método multivariado negligenciado", na década de 1970.

A variedade e riqueza da AC pode ser atribuída a diversos fatores. Em primeiro lugar, foi redescoberta e desenvolvida de várias formas, por diversos investigadores e em vários países, possuindo uma história marcada pela riqueza e variedade cultural (ver Nishisato (1994) para uma visão mais abrangente). Em segundo lugar, a sua poderosa capacidade de visualização, baseada num simples e familiar conceito geométrico, encontra aplicações nas escalas de traços psicológicos, gradientes ambientais em ecologia e ordenação de tempo em arqueologia, entre outros. Em terceiro lugar, mas não menos importante, o método pode ser aplicado a diferentes tipos de dados categóricos: contagens, preferências, avaliações, variáveis *dummy* (0/1), tornando-o numa técnica versátil em muitas áreas de investigação. (Benzécri, 1973) publicou um trabalho denominado de "L'Analyse des Données". Embora as origens teóricas desta metodologia remontem há 50 anos atrás, o verdadeiro impulso para a aplicação moderna da análise de correspondências deve-se, efetivamente, ao linguista e analista de dados francês Jean-Paul Benzécri e aos seus colegas e alunos. Inicialmente, na década de 60, trabalhou na Universidade de Rennes e, posteriormente, no campus de Jussieu, da Universidade de Paris. Desenvolvimentos paralelos sobre esta ferramenta estatística ocorreram na Holanda e Japão, centrados em torno de investigadores pioneiros como Jan de Leeuw e Chikio Hayashi. No entanto, devido à sua linguagem matemática peculiar, que torna a sua leitura difícil para os iniciados, e à pouca difusão da língua francesa, na qual a obra foi escrita, o trabalho de Benzécri não obteve a projeção desejada. Um ano mais tarde, Hill publicou um artigo no *Journal*

of the Royal Statistical Society, cujo título foi "Correspondence Analysis: A Neglected Multivariate Method" (Hill, 1974), que despertou a atenção dos países de língua inglesa para a existência desta técnica. Numa revisão bibliográfica, de publicações sobre análise de correspondências, Beh (2004) refere um aumento exponencial de artigos sobre o assunto. No entanto, esta bibliografia é focada, principalmente, em revistas metodológicas, pelo que não incluiu a explosão de aplicações de análise de correspondências em campos tão diversos como a arqueologia, linguística, pesquisa de marketing, psicologia, sociologia, educação, geologia, ecologia e medicina. Na verdade, em todas as ciências físicas, sociais, humanas e biológicas.

Embora difícil de ler para os iniciados na área, como já fora referido, o livro de Benzécri (1973) é uma fonte rica de resultados básicos e aplicações de AC. Uma simples introdução sobre AC é dada por Greenacre (1993) sob a forma estruturada e uma explicação não técnica voltada especificamente para um público interessado, principalmente, em aspetos práticos. Diferentes abordagens sobre a teoria e a prática da AC podem ser encontradas em Lebart *et al.* (1984), Greenacre (1984), Gifi (1990) e Nishisato (1990). O livro editado por Greenacre and Blasius (1994) dá uma explicação completa sobre a forma simples e múltipla da AC, incluindo um guia passo-a-passo dos cálculos e muitas aplicações no contexto das ciências sociais.

A Análise de correspondência (AC) é uma técnica exploratória multivariada para a análise gráfica e numérica de quase todos os dados em forma de matriz com entradas não negativas, mas envolve principalmente as tabelas de frequências ou contagens. A AC, basicamente, converte uma matriz de dados não negativos num tipo particular de representação gráfica em que as linhas e colunas da matriz são simultaneamente representadas numa dimensão reduzida, isto é, por pontos no gráfico.

Este método permite estudar as relações e semelhanças existentes entre:

- a) as categorias de linhas e entre as categorias de colunas de uma tabela de contingência
- b) o conjunto de categorias de linhas e o conjunto categorias de colunas. A AC mostra como as variáveis dispostas em linhas e colunas estão relacionadas e não somente se a relação existe. Embora seja considerada uma técnica descritiva e exploratória, a AC simplifica dados complexos e produz análises exaustivas de informações que suportam conclusões a respeito das mesmas.

A AC possui diversos aspetos que a distinguem de outras técnicas de análise de dados. A sua natureza multivariada permite revelar relações que não seriam detetadas com comparações de pares de variáveis. É altamente flexível quanto a pressuposições sobre os

dados: o único requisito é a existência de uma matriz retangular com entradas não negativas. Esta ferramenta pode ser estendida para analisar a presença/ausência de dados, rankings e preferências, dados de comparação pareada, de resposta múltipla, tabelas multiponto, e as tabelas de transição quadrados, entre outros. Uma vez que é orientada para dados categóricos, pode ser utilizada para analisar quase qualquer tipo de dados, após a transformação adequada, ou recodificação, como exemplificado no livro de Murtagh (2005).

5.2 Análise de Correspondências

Segundo Greenacre (1981), a análise de correspondências é especialmente indicada para descrever matrizes com grande volume de dados discretos e sem uma estrutura claramente definida *a priori*. Este método permite a visualização das relações mais importantes entre um grande conjunto de variáveis. Os resultados são apresentados sob a forma de gráficos, onde estão representadas as categorias de cada variável e onde se podem observar as relações entre estas, através das distâncias entre os pontos do gráfico (Greenacre, 1984).

Existem diferentes formas, mas matematicamente equivalentes, de definir AC. A abordagem que se seguirá é essencialmente gráfica. Em Benzécri (1973) a AC surge como uma adaptação para dados categóricos de Análise em Componentes Principais (ACP), que é um método para a identificação de dimensões explicando a maior percentagem possível de variância dos dados. Ambos os métodos são baseados em decomposições de matrizes centradas e normalizadas, usando a decomposição valor próprio-vetor próprio de uma matriz quadrada simétrica ou a decomposição de valor singular (DVS) de matrizes retangulares. Greenacre (1984), apresenta a teoria baseada na DVS, tratando-se de uma abordagem mais adequada para mostrar a relação entre as soluções de linha e as soluções de coluna. Semelhante à ACP, a AC fornece valores próprios que são valores singulares ao quadrado (denominados de inércias, em AC), percentagens de variância explicada (percentagens de inércia), cargas fatoriais (correlações com os eixos principais) e comunicações (percentagens de inércia explicada por linhas ou colunas individuais). Na ACP, são também interpretadas as representações gráficas, porém esta interpretação é menos comum do que no âmbito da AC, onde a representação gráfica é fundamental para a interpretação.

Desta forma, a AC pode ser considerada como um caso especial da ACP, porém dirigida a dados categóricos organizados em tabelas de contingência e não a dados contínuos.

5.2.1 Matriz Inicial de Dados ou Matriz de Partida

Seja \mathbf{K} uma matriz com entradas não negativas, denominada de **matriz inicial de dados** ou **matriz de partida**, cuja soma das linhas e colunas é superior a zero.

Apesar de, na AC haver uma simetria perfeita entre o estatuto das linhas e colunas da matriz, ir-se-á reservar as linhas da matriz para os indivíduos e as categorias ou modalidades de uma variável para as colunas.

Desta forma, a matriz inicial de dados é definida por

$$\mathbf{K}_{(n \times p)} = \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1p} \\ k_{21} & k_{22} & \cdots & k_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1} & k_{n2} & \cdots & k_{np} \end{pmatrix},$$

onde $k_{ij} \geq 0$ corresponde ao valor do indivíduo i na modalidade j ($i = 1, \dots, n$ e $j = 1, \dots, p$).

Totais marginais de linha:

Os **totais marginais de linha** (ou **somas em linha**) são dados por

$$k_{i.} = \sum_{j=1}^p k_{ij}, \quad i = 1, \dots, n \quad (5.1)$$

que indicam o número total de observações (frequência absoluta) associadas ao indivíduo i . Por sua vez, o vetor (coluna) das somas em linha é dado por

$$\mathbf{K}\mathbf{1}_{(p \times 1)}, \quad (5.2)$$

onde $\mathbf{1}$ é um vetor de 1's.

Totais marginais de coluna:

Analogamente, os **totais marginais de coluna** (ou **somas em coluna**) são dados por

$$k_{.j} = \sum_{i=1}^n k_{ij}, \quad j = 1, \dots, p \quad (5.3)$$

que indicam o número total de observações associadas à modalidade j . O vetor (linha) das somas em coluna é dado pelo produto matricial

$$\mathbf{1}_{(n \times 1)}^T \mathbf{K}, \quad (5.4)$$

onde $\mathbf{1}$ é, igualmente, um vetor de 1's de dimensão $(n \times 1)$.

Total de ocorrências observadas:

O **total de ocorrências observadas** ou o **número total de observações** é dado por

$$k_{..} = \sum_{i=1}^n \sum_{j=1}^p k_{ij} = \sum_{i=1}^n k_{i.} = \sum_{j=1}^p k_{.j}, \quad i = 1, \dots, n, \quad j = 1, \dots, p \quad (5.5)$$

A equação 5.5 pode escrever-se, matricialmente, pelo produto

$$k_{..} = \mathbf{1}_{(n \times 1)}^T \mathbf{K} \mathbf{1}_{(p \times 1)}, \quad (5.6)$$

5.2.2 Matriz de Correspondências

Ao analisar uma matriz de dados, em particular, uma tabela de contingência, faz pouco sentido olhar diretamente para os valores de cada célula, uma vez que, de forma geral, existe um diferente número de indivíduos associados a cada categoria. Deste modo, com o objetivo de analisar as diversas linhas (colunas) da matriz com base nas colunas (linhas), é necessário recorrer às frequências relativas.

Surge assim, uma noção fulcral na AC, a **matriz de correspondências**, cujos elementos correspondem ao quociente entre o respetivo elemento da matriz de partida e o número total de ocorrências, $k_{..}$. Desta forma, a matriz de correspondências \mathbf{F} é dada por

$$\mathbf{F} = \left(\frac{1}{k_{..}} \right) \mathbf{K}, \quad (5.7)$$

em que o termo geral é dado por

$$f_{ij} = \frac{k_{ij}}{k_{..}} \quad (5.8)$$

A matriz de correspondências (ou matriz de frequências relativas) é dada por

$$\mathbf{F}_{(n \times p)} = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1p} \\ f_{21} & f_{22} & \cdots & f_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{np} \end{pmatrix}, \quad (5.9)$$

onde $f_{ij} \in [0, 1]$, com $i = 1, \dots, n$, $j = 1, \dots, p$ é a frequência relativa do indivíduo i na categoria j .

Totais marginais de linha:

Os **totais marginais de linha**, da matriz \mathbf{F} , são dados por

$$f_{i.} = \sum_{j=1}^p f_{ij} = \sum_{j=1}^p \frac{k_{ij}}{k_{..}} = \frac{k_{i.}}{k_{..}} \quad (5.10)$$

O vetor (coluna) das somas de linhas é dado por

$$\mathbf{F}\mathbf{1}_{(p \times 1)}, \quad (5.11)$$

onde $\mathbf{1}$ é um vetor de 1's.

Totais marginais de coluna:

Os **totais marginais de coluna**, da matriz \mathbf{F} , são dados por

$$f_{.j} = \sum_{i=1}^n f_{ij} = \sum_{i=1}^n \frac{k_{ij}}{k_{..}} = \frac{k_{.j}}{k_{..}} \quad (5.12)$$

Assim, o vetor (linha) das somas de coluna é dado por

$$\mathbf{1}_{(n \times 1)}^T \mathbf{F}, \quad (5.13)$$

onde $\mathbf{1}$ é, igualmente, um vetor de 1's de dimensão $(n \times 1)$.

Total de ocorrências observadas:

O **total de ocorrências observadas** é dado por

$$f_{..} = \sum_{i=1}^n \sum_{j=1}^p f_{ij} = \sum_{i=1}^n f_{i.} = \sum_{j=1}^p f_{.j} = 1, \quad i = 1, \dots, n, \quad j = 1, \dots, p \quad (5.14)$$

Facilmente se verifica a equação acima apresentada, com base na equação (5.5), já que

$$f_{..} = \sum_{i=1}^n \sum_{j=1}^p f_{ij} = \sum_{i=1}^n \sum_{j=1}^p \frac{k_{ij}}{k_{..}} = \frac{1}{k_{..}} \underbrace{\sum_{i=1}^n \sum_{j=1}^p k_{ij}}_{k_{..}} = \frac{k_{..}}{k_{..}} = 1 \quad (5.15)$$

A equação (5.14) pode escrever-se, matricialmente, por

$$f_{..} = \mathbf{1}_{(n \times 1)}^T \mathbf{F} \mathbf{1}_{(p \times 1)} = 1, \quad (5.16)$$

Frequência relativa de linha:

A **frequência relativa da linha i** é dada por

$$r_{i.} = \frac{f_{i.}}{f_{..}} \quad i = 1, \dots, n \quad (5.17)$$

O vetor (coluna) das frequências relativas de linha, \mathbf{r} , é dada por

$$\mathbf{r} = \frac{1}{f_{..}} \mathbf{F} \mathbf{1}_{(p \times 1)} = \begin{pmatrix} f_{1.} \\ f_{2.} \\ \vdots \\ f_{n.} \end{pmatrix}_{(n \times 1)} \quad (5.18)$$

Uma vez que, pela equação 5.14, $f_{..} = 1$, a equação (5.18) pode-se escrever como

$$\mathbf{r} = \mathbf{F}\mathbf{1}_{(p \times 1)} \quad (5.19)$$

Frequência relativa de coluna:

A frequência relativa da coluna j é dada por

$$c_{.j} = \frac{f_{.j}}{f_{..}} \quad j = 1, \dots, p \quad (5.20)$$

De forma homóloga, o vetor (coluna) das frequências relativas de coluna, \mathbf{c} , é dado por

$$\mathbf{c} = \frac{1}{f_{..}} \mathbf{F}^T \mathbf{1}_{(n \times 1)} = \mathbf{F}^T \mathbf{1}_{(n \times 1)} = \begin{pmatrix} f_{.1} \\ f_{.2} \\ \vdots \\ f_{.n} \end{pmatrix}_{(p \times 1)} \quad (5.21)$$

De notar que, sendo \mathbf{K} uma matriz de partida, a matriz \mathbf{F} pode ser vista como uma estimativa da distribuição de probabilidade bivariada associada ao problema sob estudo. De igual forma, os vetores \mathbf{r} e \mathbf{c} são estimativas das distribuições de probabilidade marginais, associadas, respetivamente, aos indivíduos e às categorias das variáveis.

De seguida, apresentar-se-ão alguns conceitos fundamentais na AC e os quais se encontram ilustrados nas figuras 5.1 e 5.2.

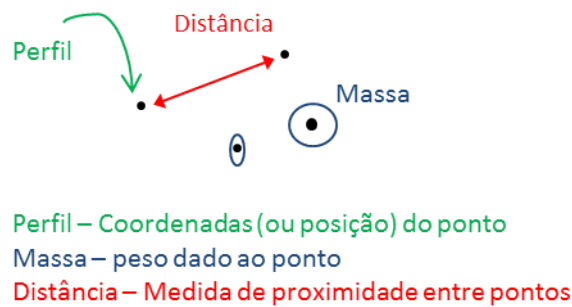


Figura 5.1: Três conceitos geométricos básicos em Análise de Correspondências.

5.2.3 Massas ou Pesos

Outro conceito fundamental na AC é o conceito de **massa** ou **peso**.

Considere-se a matriz \mathbf{F} apresentada em (5.9). A **massa** (ou **peso**) de cada um dos indivíduos é dada pela correspondente marginal de linha.

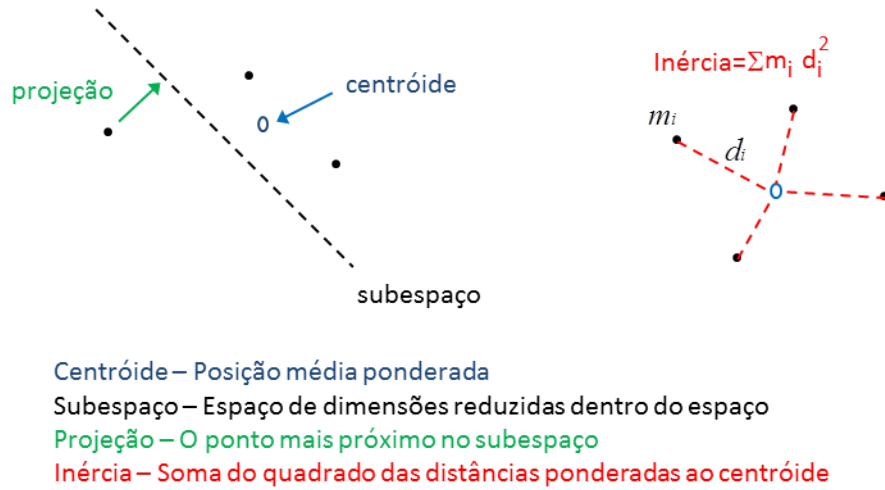


Figura 5.2: Quatro conceitos geométricos derivados em Análise de Correspondências.

Em particular, a massa do indivíduo i (peso da linha i) é dada por

$$f_{i.} = \frac{k_{i.}}{k_{..}} \quad (5.22)$$

Geralmente, define-se a matriz diagonal das massas de linha como sendo a matriz D_r dada por

$$D_r = \text{diag}(r) = \begin{pmatrix} f_{1.} & 0 & \cdots & 0 \\ 0 & f_{2.} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_{n.} \end{pmatrix}_{(n \times n)}, \quad (5.23)$$

Analogamente, a massa (ou peso) de cada uma das modalidades/categorias das variáveis é dada pelos totais das frequências marginais das colunas.

Desta forma, a massa da categoria j (peso da coluna j) é dada por

$$f_{.j} = \frac{k_{.j}}{k_{..}} \quad (5.24)$$

Define-se a matriz diagonal das massas de coluna como sendo a matriz D_c dada por

$$D_c = \text{diag}(c) = \begin{pmatrix} f_{.1} & 0 & \cdots & 0 \\ 0 & f_{.2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_{.p} \end{pmatrix}_{(p \times p)}, \quad (5.25)$$

5.2.4 Perfis

A noção de perfil surge da necessidade de eliminar a influência das marginais na comparação das linhas (colunas) da matriz.

Assim, a partir da matriz \mathbf{F} , constroem-se as matrizes denominadas de **perfil-linha** e **perfil-coluna**.

A matriz perfil-linha é tão somente resultado do quociente entre o termo de f_{ij} da linha i pela marginal da linha i , $f_{i.}$. A nova linha é denominada por perfil-linha e representada por f_j^i .

Por perfil da linha i entende-se o conjunto das frequências observadas para cada elemento dessa linha, relativas ao total de observações nessa mesma linha. Assim, a coordenada de ordem j do perfil da linha i é dada pelos p valores:

$$f_j^i = \frac{f_{ij}}{f_{i.}}, j = 1, \dots, p \quad (5.26)$$

Uma vez que $f_{ij} = \frac{k_{ij}}{k_{..}}$ (5.8) e $f_{i.} = \frac{k_{i.}}{k_{..}}$ (5.18), o perfil de linha vem dado por

$$f_j^i = \frac{f_{ij}}{f_{i.}} = \frac{\frac{k_{ij}}{k_{..}}}{\frac{k_{i.}}{k_{..}}} = \frac{k_{ij}}{k_{i.}}, j = 1, \dots, p \quad (5.27)$$

Repare-se que a soma dos p valores, isto é, a soma das coordenadas das linhas em \mathbb{R}^p , é igual a 1, uma vez que, por (5.1) e (5.27) tem-se

$$\sum_{j=1}^p f_j^i = \sum_{j=1}^p \frac{k_{ij}}{k_{i.}} = \frac{1}{k_{i.}} \underbrace{\sum_{j=1}^p k_{ij}}_{=k_{i.}} = \frac{1}{k_{i.}} k_{i.} = 1 \quad (5.28)$$

Como as coordenadas dos indivíduos têm de respeitar (5.28), estas encontram-se num espaço de dimensão $(p-1)$, por força da transformação inicial de dados específica da AC. Por exemplo, um conjunto I de indivíduos caracterizado por três categorias encontra-se no plano E da figura 5.3.

Do ponto de vista matricial, a matriz P_r dos perfis de linha calcula-se através do produto

$$P_r = D_r^{-1} \mathbf{F}, \quad (5.29)$$

onde D_r^{-1} é a matriz diagonal $(n \times n)$, cuja diagonal é dada pelos recíprocos do vetor de frequências relativas de linha, ou seja, é a matriz inversa da matriz dada em (5.23).

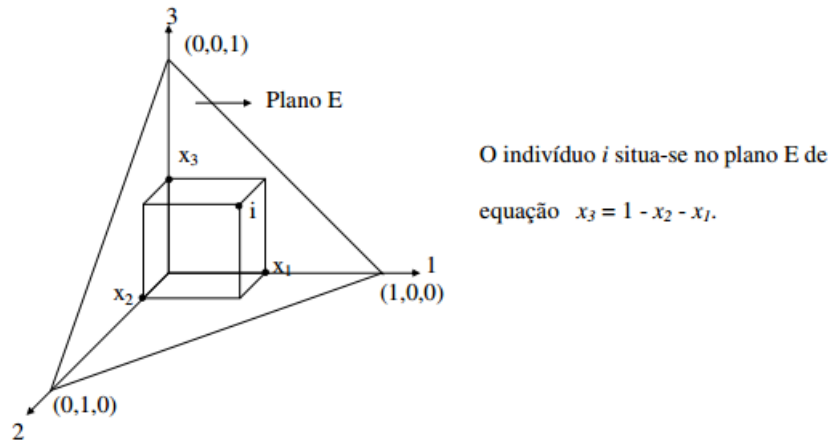


Figura 5.3: Plano que contém os indivíduos caracterizados pelas categorias 1, 2 e 3.

Assim, D_r^{-1} é definida como

$$D_r^{-1} = \begin{pmatrix} 1/f_{1.} & 0 & \cdots & 0 \\ 0 & 1/f_{2.} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/f_{n.} \end{pmatrix}_{(n \times n)} \quad (5.30)$$

Um perfil de linha corresponderá à distribuição, por variável, das observações de um indivíduo, ou seja, ao perfil do indivíduo. Desta forma, as coordenadas das linhas podem ser interpretadas como sendo a probabilidade de ocorrer j , sabendo i . Trata-se assim da probabilidade do elemento ocorrer na variável j , sabendo que se trata do indivíduo i .

De forma equivalente, o **perfil-coluna** é definido de forma semelhante. Entende-se por perfil-coluna (ou perfil de coluna) j como o conjunto das frequências observadas para cada elemento dessa coluna, relativas ao total de observações nessa coluna. Assim, a coordenada de ordem i do perfil da coluna j é dado pelos n valores:

$$f_i^j = \frac{f_{ij}}{f_{.j}}, \quad i = 1, \dots, n \quad (5.31)$$

Recorrendo às equações (5.3) e (5.31), facilmente se verifica que a soma das coordenadas de uma modalidade é, também, unitária:

$$\sum_{i=1}^n f_i^j = \sum_{i=1}^n \frac{k_{ij}}{k_{.j}} = \frac{1}{k_{.j}} \underbrace{\sum_{i=1}^n k_{ij}}_{k_{.j}} = \frac{1}{k_{.j}} k_{.j} = 1 \quad (5.32)$$

A matriz P_c dos perfis de coluna é dada, matricialmente, por

$$P_c = F D_c^{-1}, \quad (5.33)$$

onde D_c^{-1} é a matriz diagonal $(p \times p)$ cuja diagonal é dada pelos recíprocos do vetor de frequências relativas de coluna, ou seja

$$D_c^{-1} = \begin{pmatrix} 1/f_{.1} & 0 & \cdots & 0 \\ 0 & 1/f_{.2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/f_{.p} \end{pmatrix}_{(p \times p)} \quad (5.34)$$

Um perfil de coluna corresponderá à distribuição das observações de uma dada variável, por indivíduo, ou seja, ao perfil da variável. Neste caso, as coordenadas das colunas podem ser vistas como a probabilidade de acontecer i , sabendo j . Ou seja, é a probabilidade do elemento ocorrer no indivíduo i , sabendo que se trata da modalidade j .

De acordo com a notação matricial apresentada em (5.29) e (5.33), as matrizes perfil-linha e perfil-coluna que, sob o ponto de vista probabilístico, são simplesmente as probabilidades condicionais das variáveis e dos indivíduos, são dadas, respetivamente por

$$\begin{aligned} P_r = D_r^{-1}F &= \begin{pmatrix} 1/f_{1.} & 0 & \cdots & 0 \\ 0 & 1/f_{2.} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/f_{n.} \end{pmatrix} \cdot \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1p} \\ f_{21} & f_{22} & \cdots & f_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{np} \end{pmatrix} = \\ &= \begin{pmatrix} f_{11}/f_{1.} & f_{12}/f_{1.} & \cdots & f_{1p}/f_{1.} \\ f_{21}/f_{2.} & f_{22}/f_{2.} & \cdots & f_{2p}/f_{2.} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}/f_{n.} & f_{n2}/f_{n.} & \cdots & f_{np}/f_{n.} \end{pmatrix} = \begin{pmatrix} f_1^1 & f_2^1 & \cdots & f_j^1 & \cdots & f_p^1 \\ f_1^2 & f_2^2 & \cdots & f_j^2 & \cdots & f_p^2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ f_1^i & f_2^i & \cdots & f_j^i & \cdots & f_p^i \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ f_1^n & f_2^n & \cdots & f_j^n & \cdots & f_p^n \end{pmatrix}_{(n \times p)} \end{aligned} \quad (5.35)$$

$$P_c = F D_c^{-1} = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1p} \\ f_{21} & f_{22} & \cdots & f_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{np} \end{pmatrix} \cdot \begin{pmatrix} 1/f_{.1} & 0 & \cdots & 0 \\ 0 & 1/f_{.2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/f_{.p} \end{pmatrix} =$$

$$= \begin{pmatrix} f_{11}/f_{.1} & f_{12}/f_{.2} & \cdots & f_{1p}/f_{.p} \\ f_{21}/f_{.1} & f_{22}/f_{.2} & \cdots & f_{2p}/f_{.p} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}/f_{.1} & f_{n2}/f_{.2} & \cdots & f_{np}/f_{.p} \end{pmatrix} = \begin{pmatrix} f_1^1 & f_1^2 & \cdots & f_1^j & \cdots & f_1^p \\ f_2^1 & f_2^2 & \cdots & f_2^j & \cdots & f_2^p \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ f_i^1 & f_i^2 & \cdots & f_i^j & \cdots & f_i^p \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ f_n^1 & f_n^2 & \cdots & f_n^j & \cdots & f_n^p \end{pmatrix}_{(n \times p)} \quad (5.36)$$

As linhas da matriz (5.35) e as colunas da matriz 5.36 fornecem as coordenadas nos dois espaços, nomeadamente, no espaço das linhas \mathbb{R}^n e no espaço das colunas \mathbb{R}^p . De forma simplificada, cada perfil-linha (perfil-coluna) pode ser representado como um ponto no espaço onde cada elemento do perfil constitui uma coordenada. Assim, cada perfil-linha (perfil-coluna) pode ser representado como um ponto.

Desta forma, os perfis das linhas e colunas definem duas nuvens de pontos, num espaço euclidiano ponderado $(p-1)$ -dimensional e $(n-1)$ -dimensional, respetivamente.

A semelhança entre duas linhas ou duas colunas é mensurada pela distância entre os seus perfis. Quanto mais semelhantes forem os perfis, mais próximos se encontram os pontos. Por contraposição, dois perfis dissemelhantes, correspondem a pontos distantes entre si.

É através desta transformação simétrica das linhas e colunas que se assegura em AC a representação simultânea dos indivíduos e categorias no mesmo espaço.

5.2.5 Nuvens de Perfis

Cada perfil de linha, ou seja, cada linha da matriz de perfis de linha, P_r , define um ponto no espaço a p dimensões, \mathbb{R}^p . A nuvem de n pontos em \mathbb{R}^p assim resultante pode ser designada por nuvem de perfis de linha, e representada por $N(I)$.

Uma representação análoga dos perfis de coluna (colunas de P_c) é possível no espaço \mathbb{R}^n , onde se encontram as colunas da matriz. A nuvem de p pontos em \mathbb{R}^n assim resultante pode ser designada por nuvem de perfis de coluna, e representada por $N(J)$.

A nuvem de pontos $N(I)$, no espaço \mathbb{R}^p , é o conjunto dos pontos $i \in I$, cujas coordenadas são dadas pelos perfis f_j^i , com massa $f_{i.}$.

De forma semelhante, a nuvem de pontos $N(J)$, no espaço \mathbb{R}^n , é o conjunto de pontos $j \in J$, cujas coordenadas são dadas pelos perfis f_i^j , com massa $f_{.j}$.

Note-se que a menos que $p \leq 3$ e $n \leq 3$ não é possível representar estas nuvens.

Embora estas duas nuvens de pontos se encontrem em espaços diferentes ($N(I) \subset \mathbb{R}^p$ e $N(J) \subset \mathbb{R}^n$), existem várias relações entre elas.

Uma primeira relação diz respeito à dimensão dos subespaços onde cada nuvem está, efetivamente, contida. A dimensão do subespaço de \mathbb{R}^p onde está contida a nuvem dos perfis de linha, $N(I)$, é dada pela característica da matriz P_r . Tem-se assim

$$\begin{aligned} \text{car}(P_r) = \text{car}(D_r^{-1}F) &\leq \min\{\text{car}(D_r^{-1}), \text{car}(F)\} \\ &\leq \min\{n, \min(n, p)\} \\ &= \min(n, p) \end{aligned}$$

Em geral, quando não existe dependência linear nas linhas da tabela F, esta característica é precisamente a menor das dimensões da tabela.

De forma análoga, a nuvem de perfis de coluna (colunas de P_c) está num subespaço de dimensão

$$\begin{aligned} \text{car}(P_c) = \text{car}(FD_c^{-1}) &\leq \min\{\text{car}(D_c^{-1}), \text{car}(F)\} \\ &\leq \min\{p, \min(n, p)\} \\ &= \min(n, p) \end{aligned}$$

Também aqui, se não há dependência linear nas colunas da tabela F, esta característica é precisamente $\min(n, p)$, e, portanto, ambas as nuvens vivem em subespaços de igual dimensão (embora em espaços diferentes).

Uma segunda relação diz respeito aos centros de gravidade cuja definição se apresenta de seguida e a qual evidencia a relação entre as duas nuvens de pontos.

5.2.6 Centro de gravidade ou centróide

O **centro de gravidade** (ou **centróide**) de um conjunto de pontos com determinada massa é uma generalização espacial da noção de média ponderada.

O centro de gravidade dos pontos f_j^i , referentes aos indivíduos, com massa f_i , é a média ponderada da nuvem $N(I)$, que corresponde ao perfil médio.

Considerando as massas dos indivíduos, f_i , e a massa das categorias, $f_{.j}$, definidas na secção 5.2.3, pode-se determinar o centro de gravidade da nuvem dos indivíduos relativamente à categoria j , ou, por outras palavras, o centro de gravidade da nuvem das linhas

em \mathbb{R}^p , o qual é dado por

$$g_j = \sum_{i=1}^n f_i f_j^i = \sum_{i=1}^n f_i \frac{f_{ij}}{f_i} = \sum_{i=1}^n f_{ij} = \sum_{i=1}^n \frac{k_{ij}}{k_{..}} = \frac{1}{k_{..}} \sum_{i=1}^n k_{ij} = \frac{k_{.j}}{k_{..}} = f_{.j} \quad (5.37)$$

A i -ésima coordenada do centro de gravidade em \mathbb{R}^p coincide com o peso da coluna j .

Assim, a nuvem $N(I)$ de perfis de linha (que são pontos em \mathbb{R}^p) tem centro de gravidade (ponderado) dado pelo vetor das frequências relativas das colunas (pesos das categorias), apresentado em (5.21). Ou seja,

$$g_j = \mathbf{c} \quad (5.38)$$

Do mesmo modo, o centro de gravidade da nuvem das categorias relativamente ao indivíduo i é dado por

$$g_i = \sum_{j=1}^p f_{.j} f_i^j = \sum_{j=1}^p f_{.j} \frac{f_{ij}}{f_{.j}} = \sum_{j=1}^p f_{ij} = \sum_{j=1}^p \frac{k_{ij}}{k_{..}} = \frac{1}{k_{..}} \sum_{j=1}^p k_{ij} = \frac{k_{i.}}{k_{..}} = f_i. \quad (5.39)$$

A coordenada de ordem i do centro de gravidade em \mathbb{R}^n coincide com o peso da linha i .

Portanto, as coordenadas do centro de gravidade da nuvem referente às categorias, $N(J)$, são iguais às marginais de linha da matriz \mathbf{F} , que são, por conseguinte, os pesos dos indivíduos, apresentados em (5.18). Isto é,

$$g_i = \mathbf{r} \quad (5.40)$$

Assim sendo, intuitivamente tem-se que, se um perfil difere muito do perfil médio, o qual corresponde ao centro de gravidade, o seu ponto apresentar-se-á longe da origem. Por outro lado, se um perfil não diferir significativamente do perfil médio, o ponto respetivo será representado próximo do centróide.

Obviamente, caso todas as categorias possuam perfis iguais, o perfil médio coincidirá com o perfil de cada uma das categorias e, consequentemente, os pontos referentes às categorias coincidirão com o centro de gravidade.

5.2.7 Distância

A distância utilizada em AC não é a distância euclidiana habitual, mas sim uma variante da distância euclidiana, a distância euclidiana ponderada, a qual é utilizada para medir e, de certo modo, descrever, a distância entre perfis.

O centróide, sendo a média dos perfis, é também o valor esperado para cada perfil e poder-se-ia dizer que seria esperado que as linhas (colunas) tivessem esse perfil. Portanto, as distâncias entre os pontos e o centróide são distâncias entre os valores esperados e observados e, por isso, são chamadas distâncias qui-quadrado, que são semelhantes à distância euclidiana (distância métrica), mas ponderadas pelo inverso da massa (ou peso) dos perfis.

Desta forma, a distância entre dois indivíduos (linhas) i^1 e i^2 , em \mathbb{R}^p , é dada por

$$d^2(i^1, i^2) = \sum_{j=1}^p \frac{1}{f_{.j}} \left(\frac{f_{i^1j}}{f_{i^1.}} - \frac{f_{i^2j}}{f_{i^2.}} \right)^2 = \sum_{j=1}^p \frac{1}{f_{.j}} (f_j^{i^1} - f_j^{i^2})^2 \quad (5.41)$$

Analogamente, a distância entre duas categorias (colunas) j^1 e j^2 , em \mathbb{R}^n , é dada por

$$d^2(j^1, j^2) = \sum_{i=1}^n \frac{1}{f_{i.}} \left(\frac{f_{ij^1}}{f_{.j^1}} - \frac{f_{ij^2}}{f_{.j^2}} \right)^2 = \sum_{i=1}^n \frac{1}{f_{i.}} (f_i^{j^1} - f_i^{j^2})^2 \quad (5.42)$$

Esta distância, denominada de distância do χ^2 , difere da distância euclidiana (usual) pela ponderação de cada termo pelo inverso do peso e permite equilibrar a ocorrência de propriedades com diferentes ordens de grandeza, isto é, estabiliza as diferenças entre perfis, dando maior peso às proporções mais significativas.

Para além da estabilização das diferenças entre perfis, a distância do χ^2 , goza da propriedade de equivalência distribucional, que permite substituir dois indivíduos (ou duas modalidades) de perfil idêntico por um indivíduo (ou uma modalidade) de massa igual à soma das massas desses dois indivíduos (dessas duas modalidades). Por outras palavras, indivíduos (ou modalidades) idênticos (idênticas) não acrescentam informação.

As distâncias apresentadas em (5.41) e (5.42) podem ser definidas, matricialmente.

A distância entre duas linhas é definida, matricialmente, pelo produto

$$d^2(i^1, i^2) = (L^1 - L^2)^T D_c^{-1} (L^1 - L^2) ,$$

onde

$$L^1 = \begin{pmatrix} f_1^{i^1} \\ f_2^{i^1} \\ \vdots \\ f_j^{i^1} \\ \vdots \\ f_p^{i^1} \end{pmatrix}_{(p \times 1)} , \quad L^2 = \begin{pmatrix} f_1^{i^2} \\ f_2^{i^2} \\ \vdots \\ f_j^{i^2} \\ \vdots \\ f_p^{i^2} \end{pmatrix}_{(p \times 1)} \quad (5.43)$$

e D_c^{-1} é a inversa da matriz (5.34).

Assim, $d^2(i^1, i^2)$ refere-se à distância euclidiana entre i^1 e i^2 , na métrica D_c^{-1} .

Do mesmo modo, a distância entre colunas é dada, matricialmente, pelo produto

$$d^2(j^1, j^2) = (C^1 - C^2)^T D_r^{-1} (C^1 - C^2) ,$$

onde

$$L^1 = \begin{pmatrix} f_1^{j^1} \\ f_2^{j^1} \\ \vdots \\ f_i^{j^1} \\ \vdots \\ f_p^{j^1} \end{pmatrix}_{(n \times 1)} , \quad L^2 = \begin{pmatrix} f_1^{j^2} \\ f_2^{j^2} \\ \vdots \\ f_i^{j^2} \\ \vdots \\ f_p^{j^2} \end{pmatrix}_{(n \times 1)} \quad (5.44)$$

e D_r^{-1} é a matriz inversa da matriz (5.30).

Desta forma, $d^2(j^1, j^2)$ refere-se à distância euclidiana entre j^1 e j^2 , na métrica D_r^{-1} .

Note-se que a distância do χ^2 pode ser convertida numa distância euclidiana habitual, já que

$$d^2(i^1, i^2) = \sum_{j=1}^p \frac{1}{f_{.j}} \left(\frac{f_{i^1 j}}{f_{i^1 .}} - \frac{f_{i^2 j}}{f_{i^2 .}} \right)^2 = \sum_{j=1}^p \left(\frac{f_{i^1 j}}{f_{i^1 .} \sqrt{f_{.j}}} - \frac{f_{i^2 j}}{f_{i^2 .} \sqrt{f_{.j}}} \right)^2$$

Para trabalhar com a distância euclidiana usual é necessário efetuar a seguinte transformação de coordenadas:

$$\frac{f_{ij}}{f_{i.}} \longrightarrow \frac{f_{ij}}{f_{i.} \sqrt{f_{.j}}} \quad \text{e} \quad \frac{f_{ij}}{f_{.j}} \longrightarrow \frac{f_{ij}}{f_{.j} \sqrt{f_{i.}}} \quad (5.45)$$

E, portanto, as coordenadas do centro de gravidade das linhas e das colunas, dada esta transformação de coordenadas, são dadas, respetivamente, por

$$g_j = \sum_{i=1}^n f_{i.} \frac{f_{ij}}{f_{i.} \sqrt{f_{.j}}} = \sqrt{f_{.j}} \quad \text{e} \quad g_i = \sum_{j=1}^p f_{.j} \frac{f_{ij}}{f_{.j} \sqrt{f_{i.}}} = \sqrt{f_{i.}} \quad (5.46)$$

5.2.8 Inércia

Tal como na Análise de Componentes Principais (ACP), o objetivo da AC é encontrar os eixos de inércia da nuvem das linhas (ou colunas) no espaço das colunas (ou linhas) da matriz de partida. O problema é, pois, a pesquisa dos vetores próprios de uma matriz de inércia particular em que as coordenadas dos pontos são dadas pelos perfis f_j^i e f_i^j .

Uma forma de medir a dispersão da nuvem de pontos é medir a sua **inércia**, que consiste na soma de quadrados das distâncias de cada ponto, perfil-linha (perfil-coluna),

em relação ao seu centro de gravidade. A inércia surge do termo "momento de inércia", ligado à mecânica. Para um corpo de massa m , cujo centro de massa se encontra a uma distância d do centro de gravidade, o momento de inércia J é dado por $J = md^2$.

No campo da AC é possível encontrar um conceito semelhante, isto é, existe uma nuvem de pontos, com coordenadas determinadas pelos perfis, cuja massa total é 1, a qual possui um centro de gravidade (perfil médio) e onde se pode definir uma distância entre pontos, neste caso, a distância do χ^2 .

Resumidamente, tem-se:

Tabela 5.1: Resumo sobre a nuvem de pontos

Espaço \mathbb{R}^p $N(I)$	Espaço \mathbb{R}^n $N(j)$
n pontos com coordenadas $D_r^{-1}F (=P_r)$	p pontos com coordenadas $FD_c^{-1} (=P_c)$
Métrica dada por D_c^{-1}	Métrica dada por D_r^{-1}
Massa/peso r	Massa/peso c

Uma forma de medir a dispersão de cada nuvem de pontos é medir a sua inércia, ou seja, a soma de quadrados das distâncias de cada ponto, perfil de linha, em relação ao seu centro de gravidade.

O centro de gravidade da nuvem de pontos dos indivíduos é dado por $f_{.j}$ (vetor \mathbf{c}), pelo que a distância entre o perfil das linha i e o centro de gravidade, isto é, a inércia simples da nuvem de perfis de linha, é dado por

$$\sum_{j=1}^p \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - f_{.j} \right)^2 = \sum_{j=1}^p \frac{1}{f_{.j}} (f_j^i - f_{.j})^2 = \sum_{j=1}^p (f_j^i - f_{.j})^T D_c^{-1} (f_j^i - f_{.j}) \quad (5.47)$$

No entanto, existem argumentos que justificam utilizar inércias ponderadas, com métricas diferentes da euclidiana usual. De facto, o peso relativo de cada linha (ou coluna) não é igual, pelo que não faz sentido que no cálculo da inércia dos perfis de linha se dê ponderações proporcionais à frequência relativa de cada linha, ou seja, a contribuição de cada linha seja multiplicada por $f_{i.}$.

Assim, considerando a média ponderada da distância, com pesos $f_{i.}$ (vetor \mathbf{r}), obtém-se a inércia de linhas, isto é, a inércia relativa aos indivíduos:

$$\sum_{i=1}^n f_{i.} \sum_{j=1}^p (f_j^i - f_{.j})^T D_c^{-1} (f_j^i - f_{.j}) \quad (5.48)$$

Desenvolvendo a expressão (5.48) vem

$$\sum_{i=1}^n f_{i.} \sum_{j=1}^p (f_j^i - f_{.j})^T D_c^{-1} (f_j^i - f_{.j}) = \sum_{i=1}^n f_{i.} \sum_{j=1}^p \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - f_{.j} \right)^2 = \sum_{i=1}^n \sum_{j=1}^p f_{i.} \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - f_{.j} \right)^2$$

Fazendo as substituições adequadas, com recurso às equações (5.8), (5.10) e (5.12), a inércia das linhas vem dada por

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^p f_{i.} \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - f_{.j} \right)^2 = \sum_{i=1}^n \sum_{j=1}^p \frac{k_{i.}}{k_{..}} \frac{1}{\frac{k_{.j}}{k_{..}}} \left(\frac{\frac{k_{ij}}{k_{..}}}{\frac{k_{i.}}{k_{..}}} - \frac{k_{.j}}{k_{..}} \right)^2 \\ &= \sum_{i=1}^n \sum_{j=1}^p \frac{k_{i.}}{k_{..}} \frac{k_{..}}{k_{.j}} \left(\frac{k_{ij}}{k_{i.}} - \frac{k_{.j}}{k_{..}} \right)^2 = \sum_{i=1}^n \sum_{j=1}^p \frac{k_{i.}}{k_{..}} \frac{k_{..}}{k_{.j}} \frac{1}{k_{i.}^2} \left(k_{ij} - \frac{k_{i.} k_{.j}}{k_{..}} \right)^2 \\ &= \frac{1}{k_{..}} \sum_{i=1}^n \sum_{j=1}^p \frac{k_{..}}{k_{i.} k_{.j}} \left(k_{ij} - \frac{k_{i.} k_{.j}}{k_{..}} \right)^2 = \frac{1}{k_{..}} \sum_{i=1}^n \sum_{j=1}^p \frac{\left(k_{ij} - \frac{k_{i.} k_{.j}}{k_{..}} \right)^2}{\frac{k_{i.} k_{.j}}{k_{..}}} \\ &= \frac{1}{k_{..}} \chi^2 \end{aligned}$$

onde χ^2 é a estatística do qui-quadrado utilizada nos testes de independência.

Em notação matricial, facilmente se verifica que a inércia das linhas é dada por

$$In(I) = \text{traço} \left[D_r (P_r - \mathbf{1} \mathbf{c}^T) D_c^{-1} (P_r - \mathbf{1} \mathbf{c}^T)^T \right], \quad (5.49)$$

que corresponde à soma dos elementos da diagonal principal da matriz de inércia das linhas.

A inércia das colunas é, igualmente, definida como a média ponderada das distâncias dos perfis de coluna ao centro de gravidade com ponderações dadas pelos pesos $f_{.j}$ (vetor r).

$$\sum_{j=1}^p f_{.j} \sum_{i=1}^n (f_i^j - f_{i.})^T D_r^{-1} (f_i^j - f_{i.}) \quad (5.50)$$

Desenvolvendo a expressão (5.50) vem

$$\sum_{j=1}^p f_{.j} \sum_{i=1}^n (f_i^j - f_{i.})^T D_r^{-1} (f_i^j - f_{i.}) = \sum_{j=1}^p f_{.j} \sum_{i=1}^n \frac{1}{f_{i.}} \left(\frac{f_{ij}}{f_{.j}} - f_{i.} \right)^2 = \sum_{j=1}^p \sum_{i=1}^n f_{.j} \frac{1}{f_{i.}} \left(\frac{f_{ij}}{f_{.j}} - f_{i.} \right)^2$$

Fazendo as substituições adequadas, com recurso às equações (5.8), (5.10) e (5.12), a

inércia das colunas vem dada por

$$\begin{aligned}
 & \sum_{j=1}^p \sum_{i=1}^n f_{.j} \frac{1}{f_{i.}} \left(\frac{f_{ij}}{f_{.j}} - f_{i.} \right)^2 = \sum_{j=1}^p \sum_{i=1}^n \frac{k_{.j}}{k_{..}} \frac{1}{\frac{k_{i.}}{k_{..}}} \left(\frac{\frac{k_{ij}}{k_{..}}}{\frac{k_{.j}}{k_{..}}} - \frac{k_{i.}}{k_{..}} \right)^2 \\
 &= \sum_{j=1}^p \sum_{i=1}^n \frac{k_{.j}}{k_{..}} \frac{k_{..}}{k_{i.}} \left(\frac{k_{ij}}{k_{.j}} - \frac{k_{i.}}{k_{..}} \right)^2 = \sum_{j=1}^p \sum_{i=1}^n \frac{k_{.j}}{k_{..}} \frac{k_{..}}{k_{i.}} \frac{1}{k_{.j}^2} \left(k_{ij} - \frac{k_{.j} k_{i.}}{k_{..}} \right)^2 \\
 &= \frac{1}{k_{..}} \sum_{j=1}^p \sum_{i=1}^n \frac{k_{..}}{k_{.j} k_{i.}} \left(k_{ij} - \frac{k_{.j} k_{i.}}{k_{..}} \right)^2 = \frac{1}{k_{..}} \sum_{j=1}^p \sum_{i=1}^n \frac{\left(k_{ij} - \frac{k_{.j} k_{i.}}{k_{..}} \right)^2}{\frac{k_{.j} k_{i.}}{k_{..}}} \\
 &= \frac{1}{k_{..}} \chi^2
 \end{aligned}$$

Matricialmente, a inércia das colunas é dada por

$$In(J) = \text{traço} \left[D_c (P_c - \mathbf{1r}^T) D_r^{-1} (P_c - \mathbf{1r}^T)^T \right], \quad (5.51)$$

que corresponde à soma dos elementos da diagonal principal da matriz de inércia das colunas.

Tendo em conta as expressões definidas anteriormente, para a inércia das linhas e das colunas, verifica-se que a inércia das duas nuvens é igual e pode ser expressa por

$$In(I) = In(J) = \frac{1}{k_{..}} \chi^2 \quad (5.52)$$

Assim, a inércia das linhas e das colunas é igual à inércia total, ou simplesmente, inércia.

Pode-se ainda deduzir que a inércia total é igual ao quadrado do coeficiente de contingência em média quadrática, dado por

$$\Phi = \sqrt{\left(\frac{\chi^2}{k_{..}} \right)} \quad (5.53)$$

Ou seja, $\Phi^2 = \frac{\chi^2}{k_{..}} = \frac{1}{k_{..}} \chi^2$.

5.2.9 Matriz de Covariâncias

Na secção 5.2.7, referiu-se que a distância do χ^2 pode ser convertida na distância euclidiana habitual, através da transformação de coordenadas (5.45) e, consequentemente,

as coordenadas do centro de gravidade das linhas e colunas são dadas, respetivamente, pelas seguintes equações:

$$g_j = \sum_{i=1}^n f_{i.} \frac{f_{i^1 j}}{f_{i^1.} \sqrt{f_{.j}}} = \sqrt{f_{.j}} \quad \text{e} \quad g_i = \sum_{j=1}^p f_{.j} \frac{f_{i^2 j}}{f_{i^2.} \sqrt{f_{.j}}} = \sqrt{f_{i^2.}} \quad (5.54)$$

Esta transformação é fundamental para o cálculo da matriz de covariâncias, uma vez que a obtenção da mesma não é facilmente deduzível a partir da distância do qui-quadrado.

Dada a transformação da métrica apresentada em (5.45), em \mathbb{R}^p , o elemento genérico da matriz de inércia V ($p \times p$) escreve-se como produto dos pesos pelo quadrado das distâncias ao centro de gravidade.

Usando a transformação exigida pelo χ^2 , vem

$$v_{j^1 j^2} = \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1}}{\sqrt{f_{.j^1} f_{i.}}} - \sqrt{f_{.j^1}} \right) \left(\frac{f_{ij^2}}{\sqrt{f_{.j^2} f_{i.}}} - \sqrt{f_{.j^2}} \right) \quad (5.55)$$

A matriz V é, usualmente, designada como matriz de covariâncias dos indivíduos e corresponde à matriz de inércias relativas à nuvem de pontos $N(I)$.

De forma análoga, em \mathbb{R}^n , define-se a matriz W , cujo elemento genérico é dado por

$$w_{i^1 i^2} = \sum_{j=1}^p f_{.j} \left(\frac{f_{ij^1}}{\sqrt{f_{i^1.} f_{.j}}} - \sqrt{f_{i^1.}} \right) \left(\frac{f_{ij^2}}{\sqrt{f_{i^2.} f_{.j}}} - \sqrt{f_{i^2.}} \right) \quad (5.56)$$

A matriz W é, geralmente, designada como matriz de covariâncias das categorias e corresponde à matriz de inércias relativas à nuvem de pontos $N(J)$.

Após determinadas as matrizes V e W , é possível proceder à diagonalização, de forma a obter os valores próprios e os vetores próprios, que determinam os eixos principais de inércias. A matriz a diagonalizar para encontrar os valores próprios e vetores próprios (eixos principais de inércia) pode ser V ou W consoante $p < n$ ou $p > n$, respetivamente. Nos eixos principais de inércia são depois projetados os indivíduos e categorias.

5.2.10 Projeção de Indivíduos e Categorias. Princípio Baricêntrico

A matriz V de termo geral dado por (5.55), pode ser escrita como produto de duas matrizes $X^T X$. De facto, denotando

$$X = x_{ij} = \left(\frac{f_{ij} - f_{i.} f_{.j}}{\sqrt{f_{i.} f_{.j}}} \right) \quad (5.57)$$

e, calculando o produto $\sum_{i=1}^n x_{ij^1} x_{ij^2}$, tem-se:

$$\sum_{i=1}^n x_{ij^1} x_{ij^2} = \sum_{i=1}^n \left(\frac{f_{ij^1} - f_{i.} f_{.j^1}}{\sqrt{f_{i.} f_{.j^1}}} \right) \left(\frac{f_{ij^2} - f_{i.} f_{.j^2}}{\sqrt{f_{i.} f_{.j^2}}} \right) = \sum_{i=1}^n \frac{(f_{ij^1} - f_{i.} f_{.j^1})(f_{ij^2} - f_{i.} f_{.j^2})}{f_{i.} \sqrt{f_{.j^1} f_{.j^2}}}$$

Multiplicando e dividindo por $f_{i.}$, vem

$$\begin{aligned} \sum_{i=1}^n x_{ij^1} x_{ij^2} &= \sum_{i=1}^n \frac{f_{i.}}{f_{i.}} \frac{(f_{ij^1} - f_{i.} f_{.j^1})(f_{ij^2} - f_{i.} f_{.j^2})}{f_{i.} \sqrt{f_{.j^1} f_{.j^2}}} \\ &= \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1} - f_{i.} f_{.j^1}}{f_{i.} \sqrt{f_{.j^1}}} \times \frac{f_{ij^2} - f_{i.} f_{.j^2}}{f_{i.} \sqrt{f_{.j^2}}} \right) \\ &= \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1}}{f_{i.} \sqrt{f_{.j^1}}} - \frac{f_{i.} f_{.j^1}}{f_{i.} \sqrt{f_{.j^1}}} \right) \times \left(\frac{f_{ij^2}}{f_{i.} \sqrt{f_{.j^2}}} - \frac{f_{i.} f_{.j^2}}{f_{i.} \sqrt{f_{.j^2}}} \right) \\ &= \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1}}{f_{i.} \sqrt{f_{.j^1}}} - \frac{f_{.j^1}}{\sqrt{f_{.j^1}}} \right) \times \left(\frac{f_{ij^2}}{f_{i.} \sqrt{f_{.j^2}}} - \frac{f_{.j^2}}{\sqrt{f_{.j^2}}} \right) \end{aligned}$$

Uma vez que $\frac{f_{.j}}{\sqrt{f_{.j}}} = \sqrt{f_{.j}}$, a expressão anterior pode ser simplificada, sendo dada por

$$v_{j^1 j^2} = \sum_{i=1}^n x_{ij^1} x_{ij^2} = \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1}}{f_{i.} \sqrt{f_{.j^1}}} - \sqrt{f_{.j^1}} \right) \times \left(\frac{f_{ij^2}}{f_{i.} \sqrt{f_{.j^2}}} - \sqrt{f_{.j^2}} \right) \quad (5.58)$$

que corresponde à expressão dada em (5.55).

A matriz V admite um valor próprio nulo correspondente ao vetor próprio com coordenadas $\sqrt{f_{.j}}$. De facto, um dado vetor próprio é a parte da inércia total da nuvem explicada pelo respetivo valor próprio. Então, se $\sqrt{f_{.j}}$ for um vetor próprio, a inércia por ele explicada é nula, porque se trata do centro de gravidade.

Como os vetores próprios são ortogonais, e sendo u_α um vetor próprio de V diferente de $\sqrt{f_{.j}}$, vem

$$\sum_{j=1}^p u_{\alpha j} \sqrt{f_{.j}} = 0 \quad (5.59)$$

Como $\sqrt{f_{.j}}$ é vetor próprio de V , correspondente ao valor próprio nulo, qualquer outro vetor próprio de V é também vetor próprio da matriz de inércia não centrada V^* de termo geral $v_{j^1 j^2}^*$, dado por

$$v_{j^1 j^2}^* = \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1}}{f_{i.} \sqrt{f_{.j^1}}} \right) \left(\frac{f_{ij^2}}{f_{i.} \sqrt{f_{.j^2}}} \right) = \sum_{i=1}^n \frac{f_{ij^1} f_{ij^2}}{f_{i.} \sqrt{f_{.j^1} f_{.j^2}}} \quad (5.60)$$

Tal como a matriz V , também é possível escrever a matriz V^* como produto de duas matrizes $X^{*T}X^*$. O termo geral x_{ij}^* de X^* escreve-se:

$$x_{ij}^* = \frac{f_{ij}}{\sqrt{f_{i.}f_{.j}}} \quad (5.61)$$

Com efeito, tem-se

$$\sum_{i=1}^n x_{ij^1}^* x_{ij^2}^* = \sum_{i=1}^n \frac{f_{ij^1}}{\sqrt{f_{i.}f_{.j^1}}} \frac{f_{ij^2}}{\sqrt{f_{i.}f_{.j^2}}} = \sum_{i=1}^n \frac{f_{ij^1}f_{ij^2}}{f_{i.}\sqrt{f_{.j^1}}\sqrt{f_{.j^2}}} = v_{j^1j^2}^* \quad (5.62)$$

Desta forma, V^* possui um vetor próprio trivial, correspondente ao valor próprio 1.

A projeção do indivíduo i em u_α , isto é, o vetor próprio não trivial de ordem α de V ou V^* , pode escrever-se

$$\hat{\psi}_{\alpha i} = \sum_{j=1}^p \left(\frac{f_{ij}}{f_{i.}\sqrt{f_{.j}}} \right) u_{\alpha j} \quad (5.63)$$

Se a matriz V^* tiver sido diagonalizada, obtêm-se os respetivos valores próprios e a expressão (5.63) dá a projeção dos indivíduos nos eixos fatoriais.

Um raciocínio análogo em \mathbb{R}^n , diagonalizando a matriz W e obtendo os respetivos vetores próprios v_α , leva ao cálculo das projeções das categorias.

Assim, a projeção da categoria j em v_α é dada por

$$\hat{\varphi}_{\alpha j} = \sum_{i=1}^n \left(\frac{f_{ij}}{f_{.j}\sqrt{f_{i.}}} \right) v_{\alpha i} \quad (5.64)$$

As relações de transição, depois de adaptadas ao formalismo particular da AC, permitem diagonalizar apenas uma das matrizes (a de menor dimensão).

Considerem-se, assim, as seguintes relações:

$$v_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} X u_\alpha \quad (5.65)$$

e

$$u_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} X^T v_\alpha \quad (5.66)$$

Fazendo $X = X^*$, cujo termo geral foi apresentado em (5.61), a linha i da relação matricial (5.65) é dada por

$$v_{\alpha i} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \left(\frac{f_{ij}}{\sqrt{f_{i.}f_{.j}}} \right) u_{\alpha j} \quad (5.67)$$

Multiplicando ambos os membros da igualdade (5.67) pelo quociente $\frac{\sqrt{\lambda_\alpha}}{\sqrt{f_i}}$ obtém-se

$$\begin{aligned}
 \frac{\sqrt{\lambda_\alpha} v_{\alpha_i}}{\sqrt{f_i}} &= \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_i}} \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \left(\frac{f_{ij}}{\sqrt{f_i} \sqrt{f_j}} \right) u_{\alpha_j} \\
 \Leftrightarrow \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_i}} v_{\alpha_i} &= \frac{1}{\sqrt{f_i}} \sum_{j=1}^p \left(\frac{f_{ij}}{\sqrt{f_i} \sqrt{f_j}} \right) u_{\alpha_j} \\
 \Leftrightarrow \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_i}} v_{\alpha_i} &= \sum_{j=1}^p \left(\frac{f_{ij}}{\sqrt{f_i} \sqrt{f_i} \sqrt{f_j}} \right) u_{\alpha_j} \\
 \Leftrightarrow \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_i}} v_{\alpha_i} &= \sum_{j=1}^p \left(\frac{f_{ij}}{f_i \sqrt{f_j}} \right) u_{\alpha_j} = \hat{\psi}_{\alpha_i}
 \end{aligned}$$

E, portanto,

$$\hat{\psi}_{\alpha_i} = \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_i}} v_{\alpha_i} \quad (5.68)$$

ou

$$v_{\alpha_i} = \frac{\sqrt{f_i}}{\sqrt{\lambda_\alpha}} \hat{\psi}_{\alpha_i} \quad (5.69)$$

Por sua vez, a linha j da relação (5.66) escreve-se da seguinte forma

$$u_{\alpha_j} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n \left(\frac{f_{ij}}{\sqrt{f_i} \sqrt{f_j}} \right) v_{\alpha_i} \quad (5.70)$$

Multiplicando ambos os membros da igualdade (5.70) pelo quociente $\frac{\sqrt{\lambda_\alpha}}{\sqrt{f_j}}$ obtém-se

$$\begin{aligned}
 \frac{\sqrt{\lambda_\alpha} u_{\alpha_j}}{\sqrt{f_j}} &= \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_j}} \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n \left(\frac{f_{ij}}{\sqrt{f_i} \sqrt{f_j}} \right) v_{\alpha_i} \\
 \Leftrightarrow \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_j}} u_{\alpha_j} &= \frac{1}{\sqrt{f_j}} \sum_{i=1}^n \left(\frac{f_{ij}}{\sqrt{f_i} \sqrt{f_j}} \right) v_{\alpha_i} \\
 \Leftrightarrow \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_j}} u_{\alpha_j} &= \sum_{i=1}^n \left(\frac{f_{ij}}{\sqrt{f_i} \sqrt{f_j} \sqrt{f_j}} \right) v_{\alpha_i} \\
 \Leftrightarrow \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_j}} u_{\alpha_j} &= \sum_{i=1}^n \left(\frac{f_{ij}}{f_j \sqrt{f_i}} \right) v_{\alpha_i} = \hat{\varphi}_{\alpha_j}
 \end{aligned}$$

E, portanto,

$$\hat{\varphi}_{\alpha_j} = \frac{\sqrt{\lambda_\alpha}}{\sqrt{f_j}} u_{\alpha_j} \quad (5.71)$$

ou

$$u_{\alpha_j} = \frac{\sqrt{f_j}}{\sqrt{\lambda_\alpha}} \hat{\varphi}_{\alpha_j} \quad (5.72)$$

Substituindo em (5.64), o valor de v_{α_i} dado por (5.69), tem-se

$$\hat{\varphi}_{\alpha_j} = \sum_{i=1}^n \left(\frac{f_{ij}}{f_{.j} \sqrt{f_{.i}}} \right) \left(\frac{\sqrt{f_{.i}}}{\sqrt{\lambda_{\alpha}}} \hat{\psi}_{\alpha_i} \right) \Leftrightarrow \hat{\varphi}_{\alpha_j} = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{i=1}^n \left(\frac{f_{ij}}{f_{.j}} \right) \hat{\psi}_{\alpha_i} \quad (5.73)$$

Simetricamente, substituindo em (5.63), o valor de u_{α_j} , obtido em (5.72), tem-se

$$\hat{\psi}_{\alpha_i} = \sum_{j=1}^p \left(\frac{f_{ij}}{f_{.i} \sqrt{f_{.j}}} \right) \left(\frac{\sqrt{f_{.j}}}{\sqrt{\lambda_{\alpha}}} \hat{\varphi}_{\alpha_j} \right) \Leftrightarrow \hat{\psi}_{\alpha_i} = \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{j=1}^p \left(\frac{f_{ij}}{f_{.i}} \right) \hat{\varphi}_{\alpha_j} \quad (5.74)$$

As relações (5.73) e (5.74) exprimem o **princípio baricêntrico** e significam, respectivamente, que a projeção de um indivíduo é (a menos do fator de escala $\frac{1}{\lambda_{\alpha}}$) o centro de gravidade das projeções de todas as propriedades (com pesos dados pelo perfil do indivíduo) e que a projeção de uma categoria é o centro de gravidade das projeções de todos os indivíduos (a menos do fator de escala $\frac{1}{\lambda_{\alpha}}$).

É, assim, possível dar um sentido às projeções simultâneas de indivíduos e propriedades no mesmo espaço. As projeções dos elementos de uma nuvem (por exemplo, os indivíduos em \mathbb{R}^p) são o baricentro dos elementos da outra (as categorias em \mathbb{R}^n). Para a interpretação em gráficos planos, o princípio baricêntrico assegura que um indivíduo se encontre próximo das categorias com que ele mais se associa (e vice-versa).

5.2.11 Decomposição em Valores Singulares

A Decomposição em Valores Singulares (DVS) é um resultado matemático fundamental em AC, tal como para outras técnicas de redução de dimensionalidade, tais como Análise em Componentes Principais, Análise de Correlação Canônica e Análise Discriminante Linear.

A decomposição em valores singulares generalizada (DVSG) corresponde à extensão da DVS que tem em conta as restrições impostas às linhas e colunas da matriz inicial.

Antes de apresentar a DVS, comece-se por introduzir alguns conceitos necessários, nomeadamente, matriz ortogonal e valor singular.

Definição 5.1 (Matriz ortogonal).

Seja $A_{(n \times p)}$ uma matriz real. A é uma **matriz ortogonal** se $A^T A = I_{(p \times p)}^1$.

Definição 5.2 (Valor singular).

Seja $A_{(n \times p)}$ uma matriz real com característica r . Então, os valores próprios de $A^T A$ são:

$$\lambda_1 \geq \lambda_i \geq \lambda_r,$$

¹ $I_{(p \times p)}$ corresponde à matriz identidade de dimensão $(p \times p)$

onde $\sigma_i = \sqrt{\lambda_i}$ é chamado de valor singular de A , onde $i = 1, \dots, r$

Definição 5.3 (Decomposição em valores singulares).

Dada uma matriz real $A_{(n \times p)}$ de característica r , existem matrizes $U_{(n \times r)}$ e $V_{(p \times r)}$ ortogonais, tais que

$$A = U D_\sigma V^T = \sum_{i=1}^r u_i \sigma_i v_i^T, \quad (5.75)$$

onde D_σ é uma matriz diagonal cuja diagonal principal apresenta os elementos $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$.

- (1) As colunas de U são os vetores próprios de AA^T . Por outras palavras, os vetores u_1, u_2, \dots, u_r , nas colunas de U , designam-se por vetores singulares esquerdos de A e constituem uma base ortonormada para as colunas de A ;
- (2) As colunas de V são os vetores próprios de $A^T A$, isto é, os vetores v_1, v_2, \dots, v_r , nas colunas de V , designam-se por vetores singulares direitos de A e constituem uma base ortonormada para as linhas de A ;
- (3) Os elementos da diagonal de D_σ , $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, designam-se por valores singulares de A .

Em (1) e (2), $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$ são, respetivamente, os valores próprios de AA^T e $A^T A$.

Definição 5.4 (Decomposição em valores singulares Generalizada).

Sejam Ω e Φ duas matrizes simétricas definidas positivas. Então, qualquer matriz real $A_{(n \times p)}$ de característica r pode ser decomposta em

$$A = M D_\sigma N^T = \sum_{i=1}^r m_i \sigma_i n_i^T, \quad (5.76)$$

onde M e N são ortonormalizadas com respeito a Ω e Φ , respetivamente, ou seja, $M^T \Omega M = N^T \Phi N = I$

À definição 5.4 dá-se o nome de decomposição em valores singulares generalizada de A , nas métricas Ω e Φ .

- (1) As colunas de M contêm os valores singulares generalizados esquerdos de A ;
- (2) As colunas de N contêm os valores singulares generalizados direitos de A ;
- (3) A matriz diagonal D_σ contém os valores singulares generalizados de A , que correspondem aos valores singulares obtidos pela DVS de A .

Os vetores singulares generalizados de A também representam uma base ortonormada para as linhas e colunas de A , no entanto, em espaços euclidianos pesados com pesos dados, respetivamente, pelas métricas Ω e Φ .

Facilmente se demonstra a DVSG, uma vez que, para tal, basta considerar a DVS da matriz $\Omega^{-1/2}A\Phi^{-1/2}$.

Assumindo-se a decomposição em valores singulares de Ω dada por $\Omega = WD\gamma W^T$, então $\Omega^{1/2} = WD_\gamma^{1/2}W^T$

$$\Omega^{1/2}A\Omega^{1/2} = UD_\sigma V^T ,$$

com U e V matrizes ortogonais.

Determinando as matrizes M e N da DVSG de A usando, $M = \Omega^{-1/2}U$ e $N = \Phi^{-1/2}V$, obtém-se (5.76).

Após a introdução das noções de DVS e DVSG, é possível calcular as projeções no espaço euclidiano. É possível obter as soluções da AC de duas formas: através da aproximação da matriz de correspondências e através da aproximação da matriz dos perfis de linha (ou coluna).

Embora estas duas abordagens sejam distintas, produzem o mesmo resultado, pelo que se apresentará apenas a primeira das abordagens

5.2.12 Aproximação da Matriz de Correspondências

O objetivo da AC é reduzir a dimensão para um subespaço r , de preferência $r = 2$ ou $r = 3$, de forma a facilitar a visualização gráfica, com a menor perda de informação possível.

Esta redução de dimensionalidade passa pela aproximação da matriz F por uma matriz \hat{F} de característica r que minimize

$$\text{traço} \left[D_r^{-1} \left(F - \hat{F} \right) D_c^{-1} \left(F - \hat{F} \right)^T \right] = \sum_{i=1}^n \sum_{j=1}^p \frac{\left(f_{ij} - \hat{f}_{ij} \right)^2}{f_{i.} f_{.j}} \quad (5.77)$$

A melhor aproximação da matriz F é $\hat{F} = rc^T$. Porém, esta aproximação não tem interesse prático, uma vez que apenas tem em conta os totais das linhas e colunas de F . Assim sendo, é habitual a aproximação por uma matriz centrada, dada por $F^* = F - rc^T$, em vez de F .

A solução é obtida através da DVSG da matriz $F^* = UD_\alpha V^T$, supondo $U^T D_r^{-1} U = V^T D_c^{-1} = I$.

Desta forma, a DVSG da matriz F^* é dada por

$$S = D_r^{-1/2} (F - rc^T) D_c^{-1/2} = UD_\alpha V^T = \sum_{j=1}^{p-1} u_j \alpha_j v_j^T, \quad (5.78)$$

com U e V matrizes ortogonais.

Desta forma, a matriz de característica r obtém-se por eliminação das $(p-1) - r$ parcelas de (5.78), isto é,

$$\hat{F} = \sum_{j=1}^p u_j \alpha_j v_j^T \quad (5.79)$$

Tendo em conta a definição 5.4 e respetivos resultados apresentados na secção anterior, tem-se que, os vetores das colunas de U definem um conjunto de eixos coordenados, denominado de eixos principais do espaço das linhas de \mathbb{R}^p . As coordenadas correspondentes às linhas de F^* são dadas pelas colunas da matriz

$$\Psi = D_r^{-1/2} U D_\alpha \quad (5.80)$$

Analogamente, os eixos principais do espaço das colunas de \mathbb{R}^n são definidos pelos vetores das colunas de V . As coordenadas dos pontos nestes eixos são dados pelas colunas da matriz

$$\Phi = D_c^{-1/2} V D_\alpha \quad (5.81)$$

Caso se pretenda representar as linhas e colunas da matriz num gráfico bidimensional, as coordenadas correspondem às duas primeiras colunas das matrizes (5.80) e (5.81).

Tendo em conta a dedução efetuada na secção 5.2.12, conclui-se que as linhas e colunas da matriz de correspondências centradas apresentam as seguintes relações:

$$\Psi = D_r^{-1/2} F^* \Phi D_\alpha^{-1} \quad (5.82)$$

$$\Phi = D_c^{-1/2} F^{*T} \Psi D_\alpha^{-1} \quad (5.83)$$

5.2.13 Projeção de Indivíduos e Variáveis Suplementares

Em alguns casos é necessário, após efetuar uma AC de um certo conjunto de dados (sob a forma de uma matriz $(n \times p)$), considerar elementos (linhas ou colunas) "suplementares" relativamente à matriz de partida. Os elementos têm um estatuto diferente dos

$(n \times p)$ elementos principais da matriz e não intervêm na sua fatorização, pretendendo apenas encontrar-se a posição das projeções suplementares nos eixos fatoriais que resultam da diagonalização da matriz principal.

Suponhamos, assim, que após se efetuar uma AC sobre uma matriz de partida $K_{(n \times p)}$ foi adicionado um determinado número de linhas (ou colunas) suplementares, resultando dessa adição uma matriz ampliada da matriz K .

Seja K^+ a matriz ampliada da matriz K . A matriz K^+ é dada, genericamente, por

$$K^+ = \left(\begin{array}{cccc|c} k_{11} & k_{12} & \cdots & k_{1p} & k_{1q}^+ \\ k_{21} & k_{22} & \cdots & k_{2p} & k_{2q}^+ \\ \vdots & \vdots & \ddots & \vdots & \\ k_{n1} & k_{n2} & \cdots & k_{np} & k_{nq}^+ \\ \hline k_{q1}^+ & k_{q2}^+ & \cdots & k_{qp}^+ & k_{qq}^+ \end{array} \right)$$

Figura 5.4: Matriz ampliada da matriz K .

Para efetuar a representação dos elementos suplementares, basta considerar as relações (5.73) e (5.74), que exprimem o princípio baricêntrico.

Seja, sem perda de generalidade, $\frac{f_{i+j}}{f_{i+}}$ o perfil de um indivíduo suplementar.

As suas coordenadas obtêm-se, diretamente, com base na relação (5.74), sendo dadas por:

$$\hat{\psi}_{\alpha_i} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \left(\frac{f_{i+j}}{f_{i+}} \right) \hat{\varphi}_{\alpha_j}, \quad (5.84)$$

onde $\hat{\varphi}_{\alpha_j}$ são as coordenadas das categorias relativas à análise em que o indivíduo não entrou, obtidas através da fatorização da matriz de partida $(n \times p)$.

Da mesma forma, com base na relação (5.73), obtêm-se as coordenadas de uma categoria suplementar, com perfil $\frac{f_{ij+}}{f_{.j+}}$:

$$\hat{\varphi}_{\alpha_j} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n \left(\frac{f_{ij+}}{f_{.j+}} \right) \hat{\psi}_{\alpha_i} \quad (5.85)$$

Os elementos suplementares, ao contrário dos elementos (indivíduos e categorias) ativos que contribuem para a inércia total e para a posição das dimensões, podem ser vistos como pontos sem peso, cujas posições relativas fornecem informação adicional no que diz respeito à interpretação do gráfico.

5.2.14 Número de Eixos a Reter

Durante a aplicação do método de AC, podem ocorrer dois problemas:

- A superestimação do número de eixos, isto é, reter um número de eixos maior do que o adequado;
- A subestimação do número de eixos (reter um número de eixos menor que o adequado).

A superestimação dos eixos retidos tende a produzir resultados não-parcimoniosos, baseados em construtos supérfluos, com reduzido ou inadequado poder explicativo. Do mesmo modo, a subestimação de fatores retidos resulta em perda significativa de informação.

Todavia, a escolha do número de eixos a reter para analisar um conjunto de dados não é tarefa simples, uma vez que entram em linha de conta vários fatores, nomeadamente, o número de variáveis envolvidas no estudo e o tipo de problema que se está a analisar.

A solução ideal passa por reter o menor número de dimensões possível (preferencialmente dois ou três), isto é, explicar os dados pelo menor número de parâmetros. Por outro lado, pretende-se igualmente que a solução descreva os dados tão completamente quanto possível, ou seja, a percentagem de inércia explicada deve ser elevada. Procura-se, assim, um compromisso entre uma percentagem aceitável de inércia explicada e a simplicidade da interpretação - pequeno número de eixos a interpretar.

A percentagem de inércia de um eixo representa a proporção da inércia total explicada por esse eixo. Por sua vez, a inércia total da nuvem dos eixos fatoriais é igual ao somatório dos r valores próprios, onde r é o número total de eixos fatoriais.

A inércia do eixo α é λ_α .

Desta forma,

$$p^\alpha = \frac{\lambda_\alpha}{\sum_{\alpha=1}^r \lambda_\alpha} \quad (5.86)$$

representa a contribuição relativa do eixo α para a inércia total da nuvem. Multiplicando esta quantidade por 100, obtém-se a percentagem de inércia explicada pelo eixo α .

Para encontrar um equilíbrio entre os dois pressupostos atrás referidos, foram sugeridos diversos critérios, dos quais se destacam os mais comuns:

- *Scree-Plot* dos valores próprios, que consiste na representação gráfica, no eixo das abcissas, do número de eixos e, nas ordenadas, a percentagem de inércia explicada por cada eixo. Detetar uma quebra acentuada no gráfico e excluir os eixos a partir desta;
- Manter um número suficiente de eixos de modo a explicar a maior proporção de inércia ϵ (usualmente superior a 0,50), isto é, reter os primeiros q eixos que obedeçam à seguinte inequação:

$$\frac{\sum_{\alpha=1}^q \lambda_{\alpha}}{r} \geq \epsilon$$

5.2.15 Contribuições

Obtidas as projeções dos indivíduos e das categorias nos eixos fatoriais, é útil calcular alguns índices de ajuda à interpretação. Estes índices são as **contribuições absolutas e relativas**.

6.2.15.1 Contribuições Absolutas

A contribuição absoluta permite a explicação de um eixo fatorial a partir dos indivíduos I ou das categorias J , é a contribuição do elemento para a formação do eixo α .

A inércia total explicada pelo eixo λ_{α} é o produto entre as massas f_i e o quadrado das distâncias das projeções de i em α , isto é:

$$\lambda_{\alpha} = \sum_{i=1}^n f_i \hat{\psi}_{\alpha_i}^2 \quad (5.87)$$

Desta forma, uma medida da contribuição do elemento i para a inércia total explicada pelo eixo é a relação entre a parte de inércia correspondente a i e a inércia total da nuvem transportada pelo eixo α , ou seja, a contribuição absoluta do i -ésimo indivíduo é dada por

$$C_{\alpha_i}^{abs} = \frac{f_i \hat{\psi}_{\alpha_i}^2}{\lambda_{\alpha}}, \quad (5.88)$$

$$\text{com } \sum_{i=1}^n C_{\alpha_i}^{abs} = 1.$$

Analogamente, a contribuição absoluta da j -ésima categoria é dada por

$$C_{\alpha_j}^{abs} = \frac{f_j \hat{\varphi}_{\alpha_j}^2}{\lambda_{\alpha}}, \quad (5.89)$$

$$\text{com } \sum_{j=1}^p C_{\alpha_j}^{abs} = 1.$$

6.2.15.2 Contribuições Relativas

A contribuição relativa reflete a explicação de um indivíduo i ou de uma categoria j a partir dos eixos.

A inércia, em relação ao centro de gravidade, de um ponto i de \mathbb{R}^p é dada por

$$f_i \times d^2(i, g_i), \quad (5.90)$$

onde $d^2(i, g_i)$ é o quadrado da distância do ponto i ao centro de gravidade g_i .

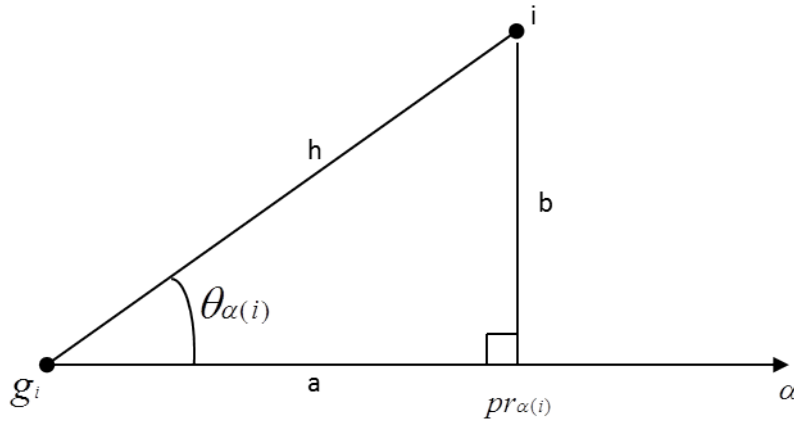


Figura 5.5: Projeção ortogonal do elemento i sobre o eixo α , $pr_{\alpha}(i)$.

Tendo em conta o Teorema de Pitágoras, ilustrado na figura 5.5, é possível escrever $d^2(i, g_i)$ como soma das coordenadas sobre esses eixos. E, portanto, tem-se

$$d_p^2(i, g_i) = \hat{\psi}_{1_i}^2 + \hat{\psi}_{2_i}^2 + \dots = \sum_{j=1}^p \left(\frac{f_{ij}}{f_i \sqrt{f_j}} - \sqrt{f_j} \right) \quad (5.91)$$

De forma análoga, o quadrado da distância de um ponto j de \mathbb{R}^n ao centro de gravidade, pode escrever-se como

$$d_n^2(j, g_j) = \hat{\varphi}_{1_j}^2 + \hat{\varphi}_{2_j}^2 + \dots = \sum_{i=1}^n \left(\frac{f_{ij}}{f_j \sqrt{f_i}} - \sqrt{f_i} \right) \quad (5.92)$$

Deste modo, com recurso a (5.91), é possível escrever a expressão geral da contribuição relativa do fator α na explicação da dispersão de um elemento da linha i , a qual se traduz por

$$C_{\alpha_i}^{rel} = \frac{\hat{\psi}_{\alpha_i}^2}{d_p^2(i, g_i)} \quad (5.93)$$

Da mesma forma, através da equação (5.92), obtém-se a seguinte expressão para a contribuição relativa do fator α na explicação da dispersão de um elemento da coluna j :

$$C_{\alpha_j}^{rel} = \frac{\hat{\varphi}_{\alpha_j}^2}{d_n^2(j, g_j)} \quad (5.94)$$

Uma vez que $\sum_{i=1}^n C_{\alpha_i}^{rel} = 1$ e $\sum_{j=1}^p C_{\alpha_j}^{rel} = 1$, é possível expressar as contribuições relativas por meio de percentagens.

A interpretação geométrica das contribuições relativas é facilmente deduzida com base na figura 5.5. Com efeito, $C_{\alpha_i}^{rel}$ é o *coseno* quadrado do ângulo θ_{α_i} formado pelo segmento g_i e pelo eixo α .

Assim, por definição de *coseno*, obtém-se a seguinte expressão:

$$\cos(\theta_{\alpha_i}) = \frac{|\hat{\psi}_{\alpha_i}|}{d_p(i, g_i)} \Leftrightarrow \cos^2(\theta_{\alpha_i}) = \frac{\hat{\psi}_{\alpha_i}^2}{d_p^2(i, g_i)} = C_{\alpha_i}^{rel} \quad (5.95)$$

Situação análoga ocorre para $C_{\alpha_j}^{rel}$.

5.2.16 Algoritmo de Análise de Correspondências

De seguida, sintetiza-se o algoritmo a usar para resolver um problema de AC:

1. Seleção do espaço de menor dimensão (seja, sem perda de generalidade, \mathbb{R}^p , isto é, admite-se que há menos categorias do que indivíduos);
2. Transformação da matriz de partida - cálculo dos perfis dos indivíduos em \mathbb{R}^p :

$$\frac{f_{ij}}{f_{i.}} = \frac{k_{ij}}{k_{i.}};$$

e das categorias, em \mathbb{R}^n :

$$\frac{f_{ij}}{f_{.j}} = \frac{k_{ij}}{k_{.j}};$$

3. Cálculo da matriz de inércia V ($p \times p$) de termo geral $v_{j^1 j^2}$:

$$v_{j^1 j^2} = \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1}}{\sqrt{f_{.j^1} f_{i.}}} - \sqrt{f_{.j^1}} \right) \left(\frac{f_{ij^2}}{\sqrt{f_{.j^2} f_{i.}}} - \sqrt{f_{.j^2}} \right)$$

ou da matriz simétrica equivalente V^* de termo geral $v_{j^1 j^2}^*$:

$$v_{j^1 j^2}^* = \sum_{i=1}^n f_{i.} \left(\frac{f_{ij^1}}{f_{i.} \sqrt{f_{.j^1}}} \right) \left(\frac{f_{ij^2}}{f_{i.} \sqrt{f_{.j^2}}} \right) = \sum_{i=1}^n \frac{f_{ij^1} f_{ij^2}}{f_{i.} \sqrt{f_{.j^1}} \sqrt{f_{.j^2}}}$$

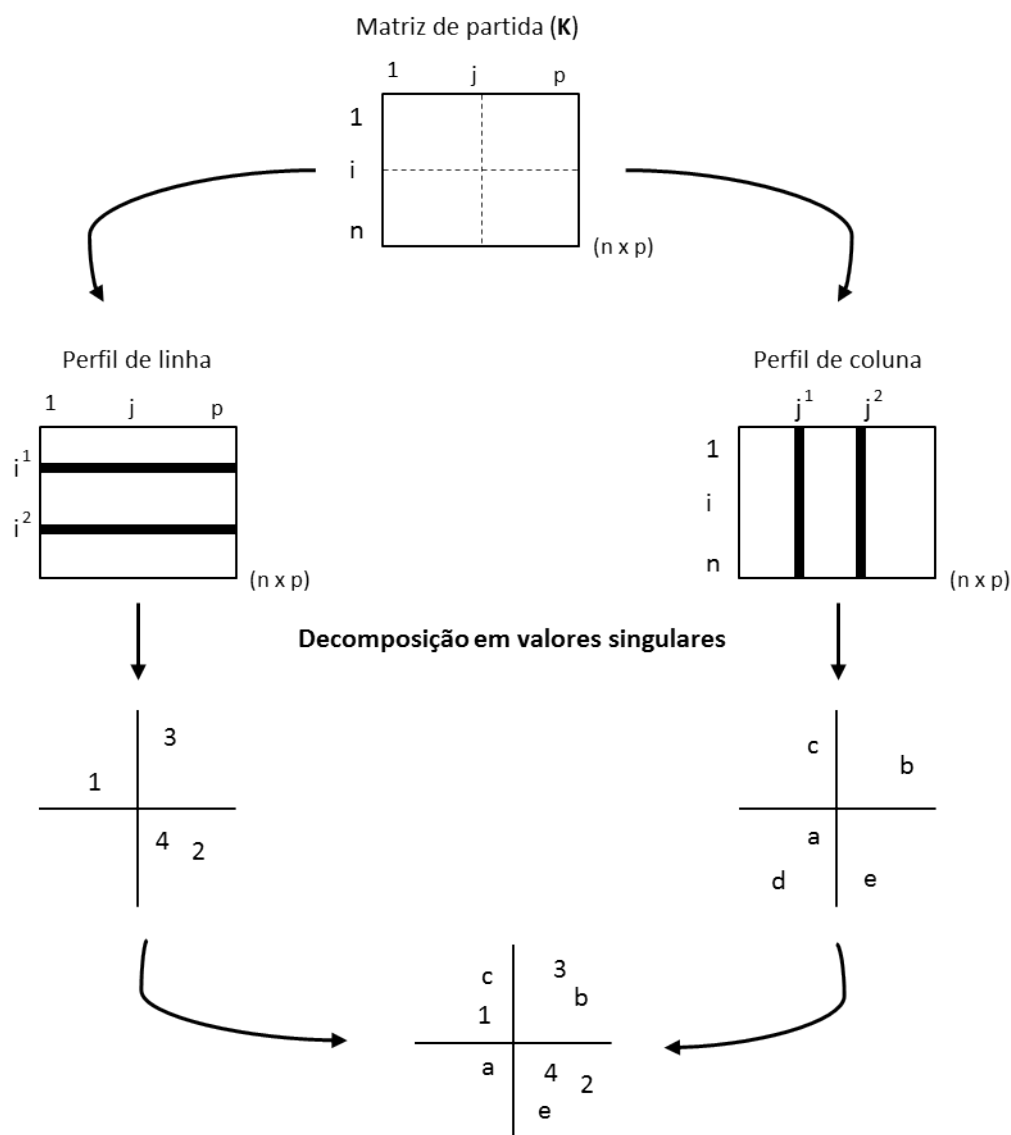


Figura 5.6: Esquema geral da AC.

4. Diagonalização da matriz de inércia V^* e obtenção dos seus valores próprios λ_α e vectores próprios u_α ;
5. Análise do *screeplot* dos valores próprios e escolha da dimensão do espaço cuja inércia acumulada explique uma percentagem significativa da inércia inicial (em geral 2 ou 3 eixos);
6. Projecção dos indivíduos nos eixos de inércia retidos:

$$\hat{\psi}_{\alpha_i} = \sum_{j=1}^p \left(\frac{f_{ij}}{f_{i.} \sqrt{f_{.j}}} \right) u_{\alpha_j}$$

7. Projeção das categorias nos eixos de inércia retidos :

$$\hat{\varphi}_{\alpha_j} = \sum_{i=1}^n \left(\frac{f_{ij}}{f_{.j} \sqrt{f_{i.}}} \right) v_{\alpha_i}$$

8. Cálculo eventual da projeção de elementos suplementares;

9. Eventual reconstituição da matriz de partida e confirmação da aproximação escolhida (número de valores próprios retidos).

10. Cálculo das contribuições absolutas e relativas;

11. Interpretação com base em valores próprios, projeção dos indivíduos e categorias nos eixos fatoriais, contribuições absolutas e relativas, etc.

5.3 Análise de Correspondências Múltiplas

5.3.1 Introdução

Na secção 5.2 apresentou-se a forma mais simples da análise de correspondências, na qual os dados são apresentados sob a forma de uma tabela resultante do cruzamento de duas variáveis (neste caso, uma referente aos indivíduos e outra relativa às categorias de uma variável), isto é, uma tabela bidimensional.

A Análise de Correspondências Múltiplas (ACM) trata da generalização da Análise de Correspondências (simples) para uma dimensão igual ou superior a três, isto é, consiste na aplicação do algoritmo de AC a matrizes de dados categóricos com mais de duas variáveis.

Nas ciências sociais, a principal aplicação da análise de correspondências múltiplas (ACM) é a visualização das inter-relações entre a variável resposta categórica e um conjunto de perguntas num questionário, como por exemplo, um conjunto de afirmações onde os entrevistados respondem de acordo com a seguinte escala: "Discordo totalmente", "Discordo", "Não concordo nem discordo", "Concordo" e "Concordo plenamente".

Considerando, por exemplo, três variáveis I , J e K , cujas categorias são indexadas por $\{i = 1, \dots, n\}$, $\{j = 1, \dots, p\}$ e $\{k = 1, \dots, q\}$, o cruzamento destas três variáveis conduz a uma tabela de contingência tridimensional, contendo npq células, como se pode visualizar na figura 5.7.

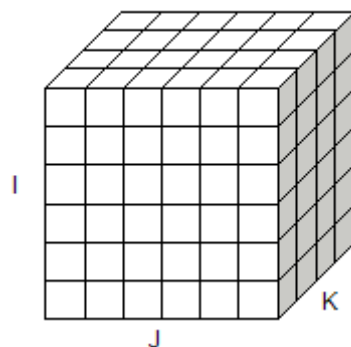


Figura 5.7: Exemplo de uma tabela tridimensional.

Antes de mais, para generalizar a Análise de Correspondências Simples, é necessário distinguir duas situações:

1. Existem várias variáveis explicativas e uma ou mais variáveis resposta, e o objeto de interesse é a relação entre estes dois conjuntos, ou seja, pretende-se com o estudo verificar qual(ais) a(s) variável(eis) que influencia(m) o comportamento da variável resposta;
2. Há um conjunto de variáveis com a mesma importância no estudo, geralmente, medidas na mesma escala, e o estudo pretende avaliar as suas inter-relações, isto é, todos os pares de relações entre as diversas variáveis na tentativa de fazer sobressair alguma característica camuflada nas várias variáveis e respetivas categorias.

A primeira situação acima referida pode ser facilmente resolvida através da concatenação das variáveis explicativas, usando o cruzamento desta informação com a variável resposta e aplicando uma AC simples sobre a tabela bidimensional obtida.

Por outro lado, para a segunda situação, o método mais usual é o de Análise de Correspondências Múltiplas o qual será desenvolvido através de duas abordagens diferentes, mas equivalentes, nas secções 5.3.3 e 5.3.4.

5.3.2 Matriz Concatenada

Esta metodologia consiste na concatenação (ou empilhamento) de matrizes, revelando-se um método de simples aplicação.

De forma a ilustrar esta técnica considere-se uma amostra de 176 indivíduos e sejam A , B , C e D as variáveis em estudo, cada uma com quatro categorias. Suponha-se, ainda,

que se pretende avaliar a relação entre a variável A (variável resposta) e cada uma das restantes variáveis (variáveis explicativas).

Seja 5.2 a tabela bidimensional que representa o cruzamento da variável A com as restantes.

Tabela 5.2: Matriz concatenada

	A1	A2	A3	A4
B1	12	8	9	10
B2	15	12	14	13
B3	7	8	16	11
B4	18	15	3	5
C1	16	14	12	9
C2	11	13	12	7
C3	16	6	9	10
C4	9	10	9	13
D1	14	17	12	8
D2	12	8	7	12
D3	8	10	11	8
D4	18	8	12	11

Após a construção da matriz concatenada, para estudar a relação entre cada uma das variáveis e a variável resposta, basta aplicar o algoritmo de AC (simples) para cada uma das três subtabelas apresentadas na tabela 5.2.

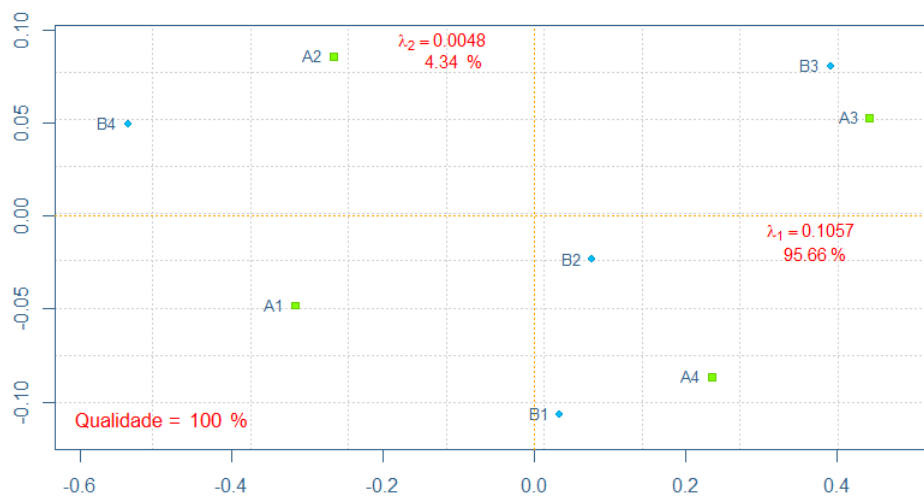


Figura 5.8: Análise de Correspondências Simples entre A e B.

```
> ac(AvsB)
```

```
[1] Resumo:
```

	Valor pp	% expl	%expl acumulada
1	0.1057	95.7	95.7
2	0.0048	4.3	100.0
3	0.0000	0.0	100.0
4	0.0000	0.0	100.0

```
[1] Linhas:
```

	Peso	CoordenadaX	CoordenadaY	CaX	CaY	CrX	CrY
B1	0.222	0.034	-0.106	0.2	52.4	9.5	90.5
B2	0.307	0.077	-0.023	1.7	3.5	91.6	8.4
B3	0.239	0.392	0.081	34.7	32.3	95.9	4.1
B4	0.233	-0.536	0.049	63.3	11.9	99.2	0.8

```
[1] Colunas:
```

	Peso	CoordenadaX	CoordenadaY	CaX	CaY	CrX	CrY
A1	0.295	-0.315	-0.048	27.8	14.3	97.7	2.3
A2	0.244	-0.265	0.086	16.2	37.2	90.6	9.4
A3	0.239	0.443	0.053	44.4	13.7	98.6	1.4
A4	0.222	0.235	-0.087	11.6	34.7	88.0	12.0

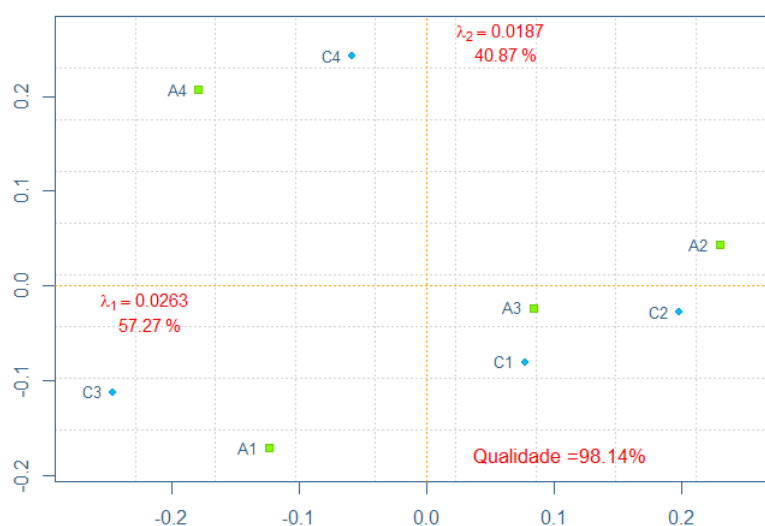


Figura 5.9: Análise de Correspondências Simples entre A e C.

```
> ac(AvsC)
```

```
[1] Resumo:
```

	Valor pp	% expl	%expl acumulada
1	0.0263	57.3	57.3
2	0.0187	40.9	98.1
3	0.0009	1.9	100.0
4	0.0000	0.0	100.0

```
[1] Linhas:
```

	Peso	CoordenadaX	CoordenadaY	CaX	CaY	CrX	CrY
C1	0.290	0.077	-0.081	6.6	10.1	47.6	52.4
C2	0.244	0.199	-0.028	36.7	1.0	98.1	1.9
C3	0.233	-0.246	-0.112	53.7	15.7	82.7	17.3
C4	0.233	-0.058	0.242	3.0	73.1	5.4	94.6

```
[1] Colunas:
```

	Peso	CoordenadaX	CoordenadaY	CaX	CaY	CrX	CrY
A1	0.295	-0.124	-0.171	17.2	46.3	34.3	65.7
A2	0.244	0.230	0.043	49.3	2.4	96.6	3.4
A3	0.239	0.084	-0.024	6.4	0.7	92.5	7.5
A4	0.222	-0.179	0.207	27.1	50.6	42.9	57.1

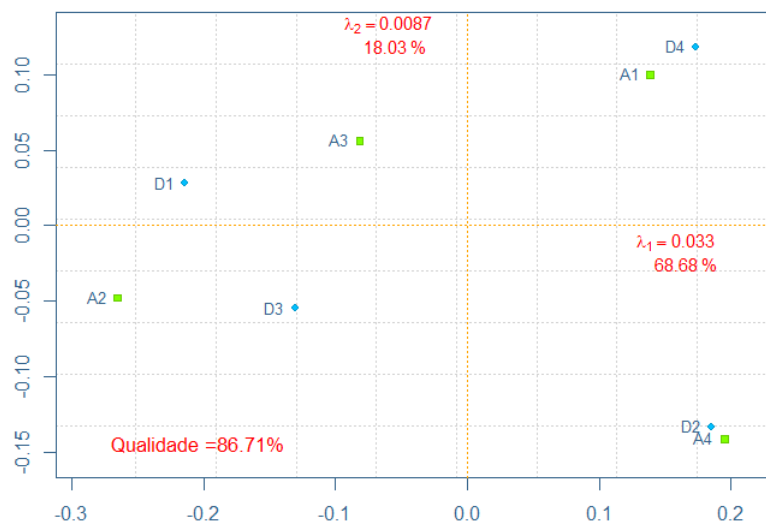


Figura 5.10: Análise de Correspondências Simples entre A e D.

```
> ac(AvsD)
[1] Resumo:
      Valor pp % expl %expl acumulada
1    0.0330   68.7      68.7
2    0.0087   18.0      86.7
3    0.0064   13.3     100.0
4    0.0000    0.0     100.0
[1] Linhas:
      Peso CoordenadaX CoordenadaY  CaX  CaY  CrX  CrY
D1 0.290      -0.214      0.028 40.3  2.6 98.3  1.7
D2 0.222       0.186     -0.133 23.2 45.3 66.1 33.9
D3 0.210      -0.131     -0.055 10.9  7.3 85.1 14.9
D4 0.278       0.174      0.118 25.5 44.8 68.5 31.5
[1] Colunas:
      Peso CoordenadaX CoordenadaY  CaX  CaY  CrX  CrY
A1 0.295       0.139      0.100 17.3 34.0 66.0 34.0
A2 0.244      -0.265     -0.048 52.2  6.4 96.9  3.1
A3 0.239      -0.082      0.056  4.8  8.7 67.8 32.2
A4 0.222       0.195     -0.141 25.6 50.9 65.7 34.3
```

Note-se que as categorias da variável resposta *A* surgem em posições distintas em cada um dos três gráficos. Facilmente se verifica que esta situação se deve ao facto da posição das categorias de uma variável dependerem da(s) variável(eis), com a a qual aquela variável é cruzada. Esta observação mostra que a posição das modalidades duma variável não é, geralmente, a mesma quando essa variável é cruzada com variáveis distintas.

Refira-se, ainda, que os gráficos 5.8, 5.9 e 5.10 não possuem qualquer significado quando o objetivo é analisar a relação entre as variáveis *B*, *C* e *D*. Com efeito, para se poder analisar a relação entre as quatro variáveis é necessário obter uma mapa do cruzamento entre cada uma das variáveis entre si.

Para responder a esta questão, existem duas soluções alternativas e, igualmente, eficazes. De salientar que se trata de soluções semelhantes à abordagem apresentada na secção 5.3.2, uma vez que consistem na conversão da matriz inicial de dados numa outra tabela que permite a representação simultânea de todas as categorias e onde é possível estabelecer a relação entre cada uma delas.

As duas tabelas ou matrizes que resultam desta transformação são denominadas de Matriz Disjuntiva Completa e Matriz de Burt e serão apresentadas nas duas secções seguintes.

5.3.3 Matriz Disjuntiva Completa

Uma das abordagens mais utilizadas na generalização da AC a tabelas multidimensionais é a aplicação do algoritmo de AC a uma matriz denominada de **matriz disjuntiva completa** (DC) ou também de matriz indicatriz.

Na matriz disjuntiva completa, as linhas representam os indivíduos e as colunas dizem respeito às categorias das variáveis em estudo. As categorias devem ser mutuamente exclusivas e exaustivas, isto é, cada indivíduo deve escolher uma, e somente uma, categoria para cada variável.

Esta matriz, ao contrário da matriz de Burt, que será abordada na secção seguinte, possui a vantagem de manter a representação das linhas da matriz de partida. Por outras palavras, as linhas da matriz DC dizem respeito aos indivíduos, à semelhança do que acontece com a matriz inicial de dados.

De seguida apresentar-se-á o método de conversão de uma matriz inicial de dados numa matriz DC:

Considerem-se Q variáveis, $q = 1, 2, \dots, Q$, onde cada uma das variáveis q possui J_q categorias, pelo que o número total de categorias das Q variáveis é dado por

$$J = \sum_{q=1}^Q J_q$$

Seja I , o conjunto dos n indivíduos em estudo, isto é, manter-se-á a notação introduzida na secção 5.2, no que diz respeito aos indivíduos.

A matriz disjuntiva completa, denote-se por Z , possui n linhas e J colunas, na qual se apresentam as respostas de cada um dos n indivíduos sob a forma binária - 0 ou 1. A tabela Z é resultado de uma justaposição de Q tabelas (ou matrizes), isto é, é da forma

$$Z_{(n \times J)} = [Z_1, Z_2 \dots, Z_q, \dots, Z_Q]$$

De forma a concretizar a ideia aqui exposta, apresenta-se, na figura 5.11, a forma genérica da matriz disjuntiva completa.

A partir da figura 5.11 pode observar-se que:

1. z_{ik} é uma variável binária ou indicatriz, isto é:

$$z_{ik} = \begin{cases} 1, & \text{se o indivíduo } i \text{ está afeto à categoria } k \\ 0, & \text{se o indivíduo } i \text{ não escolheu a categoria } k \end{cases}$$

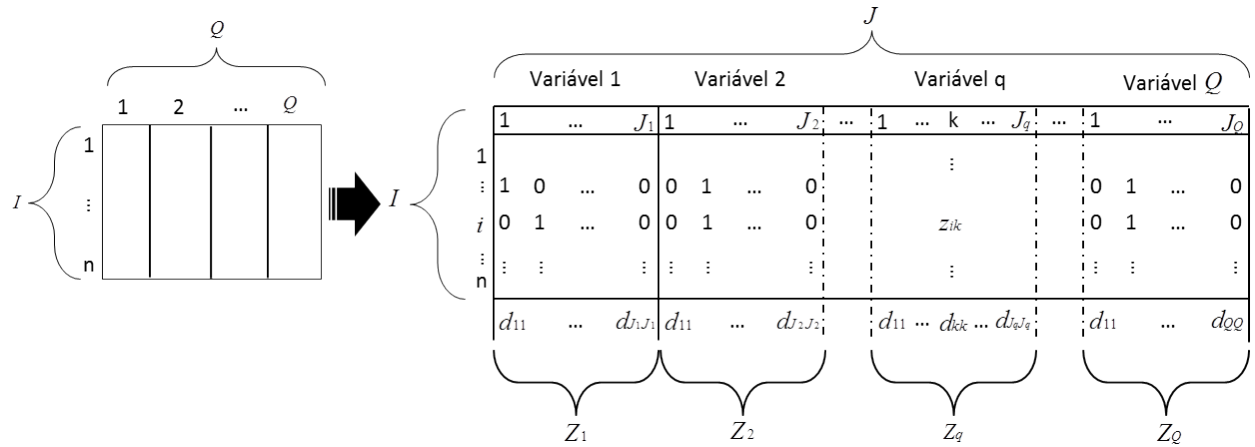


Figura 5.11: Matriz disjuntiva completa genérica.

2. Cada indivíduo deve escolher uma categoria para cada uma das Q variáveis o que, analiticamente, se traduz por:

$$\sum_{k=1}^J z_{ik} = Q, \quad i = 1, \dots, n$$

Esta propriedade da matriz DC indica simplesmente que a marginal das linhas é a mesma e corresponde ao número de variáveis;

3. Cada indivíduo deve escolher uma só categoria para cada uma das variáveis, o que é equivalente a dizer que a marginal das linhas é igual a Q (número de variáveis). Esta propriedade traduz-se pela seguinte expressão:

$$\sum_{k=1}^{J_q} z_{ik} = 1, \quad i = 1, \dots, n, \quad q = 1, \dots, Q;$$

4. A soma em coluna (ou marginal de coluna) corresponde ao número de indivíduos que responderam à categoria k :

$$\sum_{i=1}^n z_{ik} = d_{kk}, \quad k = 1, \dots, J_q$$

5. Os totais das marginais de coluna de cada uma das matrizes Z_i ($i = 1, \dots, Q$) corresponde ao número de indivíduos:

$$\sum_{k=1}^{J_q} d_{kk} = n, \quad q = 1, \dots, Q$$

Para facilitar a compreensão, na figura 5.12, apresenta-se um exemplo da construção da matriz disjuntiva completa, usando o mesmo exemplo da secção anterior. Por uma questão de simplicidade apresentam-se apenas as 10 primeiras linhas.

Uma vez que se possui quatro variáveis, cada uma com quatro categorias, tem-se que $Q = 4$, com $J_q = 4$, para $q = 1, \dots, 4$. Consequentemente, $J = \sum_{q=1}^4 J_q = 16$, pelo que o número total de categorias é 16, que corresponde ao número de colunas da matriz disjuntiva completa.

indivíduo	A	B	C	D		indivíduo	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4
I1	3	2	1	3		I1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0
I2	1	1	3	1		I2	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0
I3	3	3	4	4		I3	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1
I4	3	3	1	4		I4	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1
I5	3	1	1	4		I5	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	1
I6	4	4	2	3		I6	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0
I7	1	1	3	1		I7	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0
I8	1	3	4	4		I8	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1
I9	2	3	2	3		I9	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0
I10	1	1	1	3		I10	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0

Figura 5.12: Conversão de uma tabela para a forma disjuntiva completa.

Após a transformação da matriz de partida na matriz disjuntiva completa, da qual, na figura 5.12 se apresentam as primeiras 10 linhas, para obter a representação gráfica pretendida, basta aplicar o algoritmo da AC simples à tabela disjuntiva completa (para os 173 indivíduos)(figura 5.13).

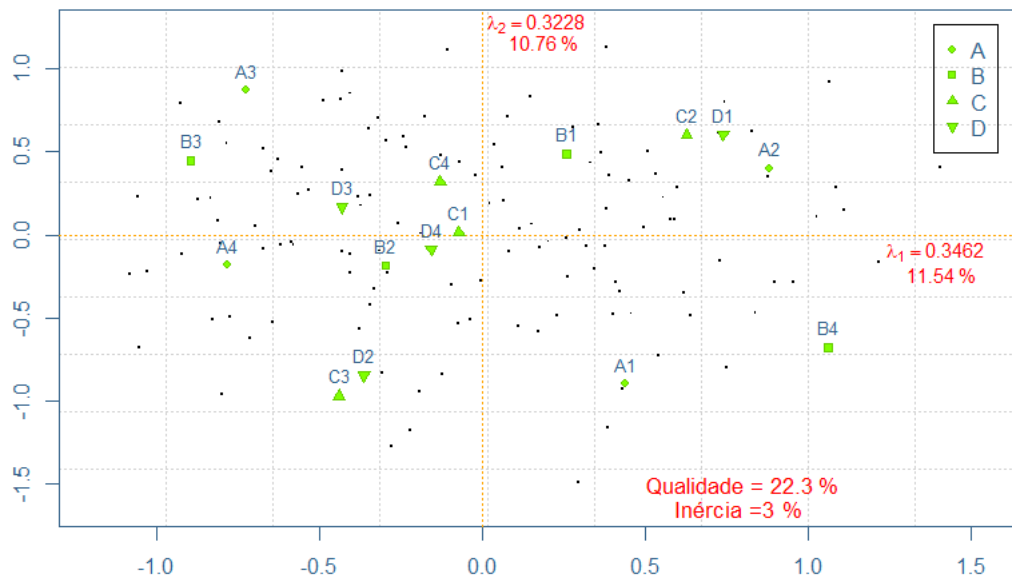


Figura 5.13: Análise de Correspondências Múltiplas via matriz disjuntiva completa.

Ao contrário do que acontece quando se considera a abordagem via matriz concatenada, as coordenadas dos pontos do mapa gerado através da AC aplicada sobre a matriz disjuntiva completa são resultado do cruzamento entre todas as variáveis.

Repare-se, também a ordem de grandeza das inércia é diferente quando considerada a ACM. Esta observação pode levar a crer que, para se obter uma representação gráfica dos dados aceitáveis, é necessário um maior número de dimensões.

Todavia, mais adiante, ir-se-á provar que esta ideia é errada, pelo que a percentagem de variância explicada de 22.3% (figura 5.13) não é desmotivante.

5.3.4 Matriz de Burt

A segunda abordagem que será apresentada nesta secção consiste na aplicação do algoritmo da AC a uma matriz denominada de **matriz de Burt**.

A matriz de Burt, denotada por B , obtém-se com base na matriz disjuntiva completa, através da seguinte relação:

$$B = Z^T Z = \begin{pmatrix} Z_1^T Z_1 & Z_1^T Z_2 & \cdots & Z_1^T Z_Q \\ Z_2^T Z_1 & Z_2^T Z_2 & \cdots & Z_2^T Z_Q \\ \vdots & \vdots & \ddots & \vdots \\ Z_Q^T Z_1 & Z_Q^T Z_2 & \cdots & Z_Q^T Z_Q \end{pmatrix}, \quad (5.96)$$

onde Z corresponde à matriz disjuntiva completa apresentada na secção anterior.

A matriz B possui as seguintes características:

1. B é uma justaposição de Q tabelas, isto é, é formada por Q submatrizes, onde Q , à semelhança do que foi referido na secção anterior, corresponde ao número de variáveis;
2. As Q submatrizes da forma $Z_q^T Z_q$ ($q = 1, \dots, Q$) são matrizes diagonais, cujos elementos (da diagonal principal) correspondem às frequências marginais de coluna da matriz Z_q . Estas submatrizes não são mais do que as tabelas de contingência, resultantes do cruzamento da variável q consigo própria;
3. As submatrizes de B da forma $Z_q^T Z_{q'}$, com $q \neq q'$, correspondem às tabelas de contingência bidimensionais resultantes do cruzamento da variável q com a variável q' ;

4. B é simétrica definida positiva, pelo que produz dois conjuntos idênticos de coordenadas para as linhas e colunas (igualdade entre linhas e colunas da matriz B). Estes conjuntos correspondem às colunas da matriz Z . Deste modo, a AC via matriz de Burt através as linhas ou colunas é equivalente.

Na figura 5.14 ilustra-se a forma genérica da matriz B .

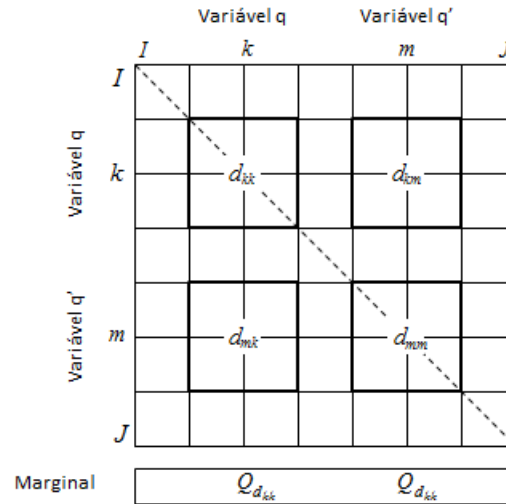


Figura 5.14: Matriz de Burt genérica.

De forma a ilustrar a construção da matriz de Burt, retome-se o exemplo das secções anteriores. Na figura 5.15 apresenta-se a conversão da matriz disjuntiva completa apresentada em 5.12 na matriz de Burt respetiva.

indivíduo	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4
I1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0
I2	1	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0
I3	0	0	1	0	0	0	1	0	0	0	1	0	0	0	1	0
I4	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0
I5	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0
I6	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0
I7	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0
I8	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1
I9	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0
I10	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0

	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4
A1	4	0	0	0	3	0	1	0	1	0	2	1	2	0	1	1
A2	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0
A3	0	0	4	0	1	1	2	0	3	0	0	1	0	0	1	3
A4	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0
B1	3	0	1	0	4	0	0	0	2	0	2	0	2	0	1	1
B2	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0
B3	1	1	2	0	0	0	4	0	1	1	0	2	0	0	1	3
B4	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0
C1	1	0	3	0	2	1	1	0	4	0	0	0	0	0	2	2
C2	0	1	0	1	0	0	1	1	0	2	0	0	0	0	2	0
C3	2	0	0	0	2	0	0	0	0	0	2	0	2	0	0	0
C4	1	0	1	0	0	0	2	0	0	0	2	0	0	0	0	2
D1	2	0	0	0	2	0	0	0	0	0	2	0	2	0	0	0
D2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D3	1	1	1	1	1	1	1	1	2	2	0	0	0	0	4	0
D4	1	0	3	0	1	0	3	0	2	0	0	2	0	0	0	4

Figura 5.15: Conversão de uma tabela disjuntiva completa na matriz de Burt.

Repare-se que, tal como fora referido, a diagonal da matriz de Burt, também designada por bloco diagonal, é composta pelas matrizes diagonais relativas a cada uma das variáveis.

Existem tantas submatrizes no bloco diagonal da matriz B quanto o número de variáveis em estudo. Neste caso, existem quatro variáveis ($Q = 4$), portanto, no bloco diagonal existem quatro matrizes diagonais, onde na diagonal das mesmas aparece o número total de ocorrências por categoria das diversas variáveis. Por exemplo, a marginal de coluna da matriz Z para a categoria $A1$ (ou, por outras palavras, a marginal de coluna da matriz Z_1 para a categoria $A1$) é quatro, isto é, quatro indivíduos estão afetos à categoria $A1$, pelo que o primeiro elemento da diagonal da primeira destas quatro matrizes é o quatro.

Para as submatrizes fora da diagonal da matriz de Burt, encontram-se as tabelas resultantes do cruzamento entre as diversas variáveis. Em cada um dos triângulos (superior e inferior) existem $Q(Q - 1)/2 = (4 \times (3 - 1))/2 = 6$ submatrizes, isto é, existem seis possíveis cruzamentos entre as variáveis em estudo.

Observando mais atentamente a figura 5.15, facilmente se percebe a poderosa capacidade de condensação de informação que a matriz B suporta. Por exemplo, pela simples observação da tabela de Burt, verifica-se que existem dois indivíduos afetos simultaneamente às categorias $B1$ e $D1$. Comparando as colunas da matriz Z , referentes às categorias $B1$ e $D1$, linha a linha, verifica-se que, das quatro vezes que ocorre $B1$, apenas duas vezes ocorre $D1$, mais especificamente, os indivíduos $I2$ e $I7$ estão afetos simultaneamente a ambas as categorias.

Aplicando o algoritmo da AC via matriz de Burt, para o mesmo conjunto de dados atrás apresentado (173 indivíduos e quatro variáveis), obtém-se o mapa apresentado em 5.16.

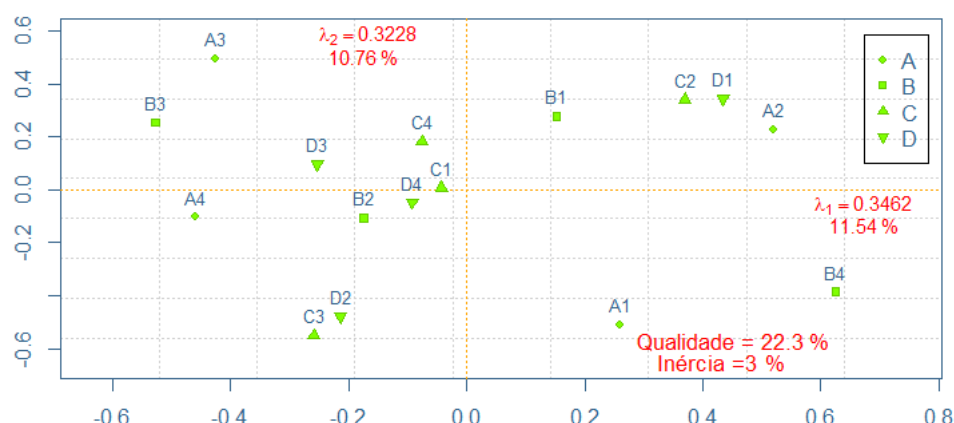


Figura 5.16: Análise de correspondências múltiplas via matriz de Burt.

Devido à relação (5.96), não é, de todo, surpreendente que as soluções obtidas em 5.13 e 5.16 se encontrem relacionadas. De facto, os gráficos são muito semelhantes, verificando-se apenas uma ligeira alteração na escala de ambos os eixos.

A única diferença entre as duas análises tem a ver com o facto de que, na versão baseada na matriz de Burt, as coordenadas principais são reduzidas em escala, enquanto que na ACM via matriz DC, a redução ocorre mais sobre o segundo eixo do que no primeiro.

As duas formas estão relacionadas como se segue:

1. Em ambas as análises, as coordenadas dos pontos das categorias são idênticos - resulta diretamente da relação (5.96);
2. Também como resultado direto da relação (5.96), as inércias da matriz de Burt correspondem ao quadrado das inércias obtidas via matriz disjuntiva completa;
3. Uma vez que as inércias são inferiores a 1, quadrar as mesmas retorna um valor mais pequeno. As coordenadas correspondem às coordenadas multiplicadas pela raiz quadrada das inércias, o que representa a redução da escala verificada aquando a comparação entre as duas representações;
4. As percentagens de inércia são, portanto, maiores quando considerada a matriz de Burt.

5.3.5 Inércias Ajustadas

Comparando as representações gráficas obtidas através da aplicação da ACM via matriz disjuntiva completa e via matriz de Burt (figuras 5.13 e 5.16), com os mapas obtidos aplicando o algoritmo de AC à matriz concatenada (figuras 5.8, 5.9 e 5.10), verifica-se que, nos dois primeiros gráficos, a percentagem de inércia associadas aos dois primeiros eixos principais é muito reduzida quando comparada com a percentagem de inércia associada aos dois primeiros eixos obtidos através da matriz concatenada.

Esta diferença na qualidade da representação, deve-se não à pior representação gráfica da ACM, mas sim às características da matriz sobre a qual é efetuada a análise. Tal situação deve-se ao facto de, na ACM, a codificação dos dados implicar a criação de variáveis na forma disjuntiva completa, criando dimensões "artificiais", isto é, cada variável é dividida em tantas variáveis quanto o número de categorias que a compõem. Como consequência, a inércia do espaço de solução ótima é artificialmente inflacionada,

originando uma subestimação nas diversas dimensões, refletindo-se, com maior impacto, nas primeiras.

Benzécri (1973) apresentou uma forma de correção, de forma a colmatar esta situação, denominada de **inércias ajustadas**.

As inércias ajustadas são obtidas através da seguinte relação:

$$\lambda_{\alpha}^* = \begin{cases} (Q/(Q-1))^2 (\lambda_{\alpha} - (1/Q)), & \text{se } \lambda_{\alpha} \geq 1/Q \\ 0, & \text{se } \lambda_{\alpha} < 1/Q \end{cases} \quad (5.97)$$

De forma a concretizar este conceito, considere-se a análise efetuada na representação gráfica 5.8, isto é, considerar-se-ão apenas as variáveis *A* e *B* e aplicar-se-á a ACM via matriz disjuntiva completa, com o intuito de comparar as soluções e as inércias de ambas as análises.

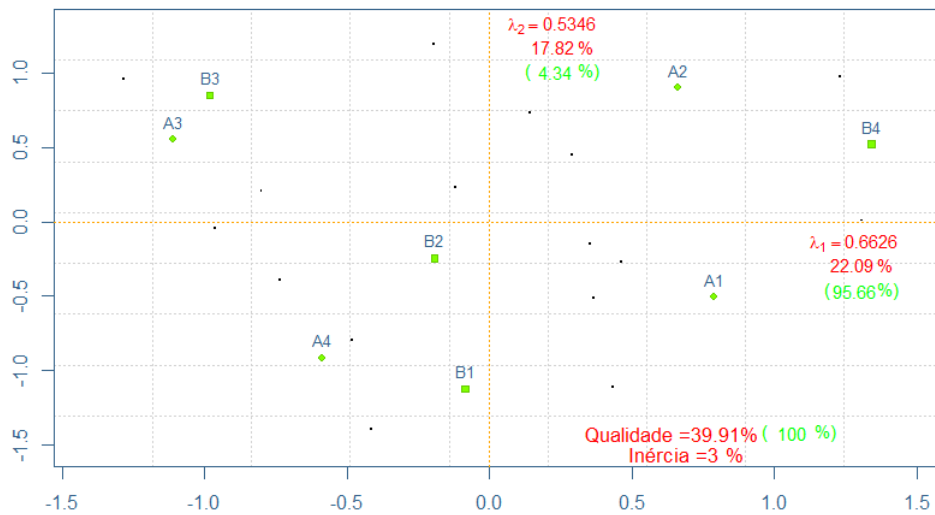


Figura 5.17: Análise de Correspondências Múltiplas via matriz disjuntiva completa *A* vs *B*.

No mapa encontram-se, a vermelho, as inércias dos dois primeiros eixos (22,09% e 17,82%, respetivamente) e a verde encontram-se as inércias ajustadas (95,66% e 4,34%, respetivamente). Relembrando o gráfico 5.8, observa-se que as inércias ajustadas obtidas através da aplicação da ACM via matriz DC, coincidem com as inércias obtidas através da aplicação do algoritmo de AC sobre a matriz concatenada.

Esta comparação gráfica, realça a consistência entre estes métodos: a menos de uma diferença na escala das categorias das variáveis, os dois gráficos são semelhantes. Em particular, as categorias perservam, espacialmente, as mesmas relações.

Através deste exemplo prático, pode-se verificar que a correção sugerida por Benzécri (1973) e demonstrada por Greenacre (2005), é fulcral na interpretação da qualidade do gráfico bidimensional, mostrando-se válida para a aplicação em estudos de maior dimensão.

5.3.6 Análise Classificatória como Auxiliar na Interpretação da ACM

A técnica de Análise Classificatória pode ser auxiliar na interpretação dos resultados da Análise de Correspondências. Esta correspondência é válida, uma vez que é possível obter-se a mesma disposição gráfica da AC no dendrograma obtido através da Análise Classificatória, o qual permite uma melhor visualização das variáveis associadas.

Na maioria das vezes em que se pretende estudar a associação das variáveis utilizando a Análise de Correspondências, em especial, a ACM, é comum depararmo-nos com grandes quantidades de variáveis e/ou categorias. Este facto pode resultar numa baixa interpretação através da solução gráfica bidimensional, o que pode requerer uma análise em dimensões muito maiores. Desta forma, torna-se necessário a utilização de uma segunda técnica de análise. A ACM (e também a AC) permite, para além de representar graficamente todas as combinações possíveis, calcular as coordenadas das categorias consideradas. Consequentemente, as coordenadas das modalidades representam dados contínuos que retratam valores de distâncias.

A Análise Classificatória é um método de agrupamento de unidades estatísticas (indivíduos, objetos, etc...), ou de variáveis, em grupos (cuja características derivam dos dados) de tal forma que objetos situados dentro do mesmo grupo são mais semelhantes do que objetos situados em grupos distintos.

O objetivo principal da Análise Classificatória consiste em classificar um conjunto de unidades estatísticas (ou variáveis) segundo um conjunto de grupos mutuamente exclusivos, exaustivos e homogêneos.

Os dados a classificar podem ser de diversos tipos:

- uma matriz de dados resumindo as medidas de um grupo de variáveis (quantitativas ou qualitativas) feitas sobre um dado conjunto de unidades estatísticas;
- uma matriz de semelhanças (proximidades) ou de dissemelhanças (distâncias);
- um conjunto de dados de preferência (resultam da ordenação de um conjunto de itens segundo algum critério ou preferência).

De acordo com o que foi referido anteriormente, ir-se-á considerar uma matriz de dissemelhanças (distâncias) que corresponde às coordenadas das categorias nas dimensões consideradas no gráfico da ACM.

Existem diversos métodos de Análise Classificatória, os quais se dividem basicamente em três tipos:

- Métodos gráficos → Partição
- Métodos Hierárquicos → Classificação Hierárquica
- Métodos Não Hierárquicos → Classificação Não Hierárquica

Para o presente trabalho foi utilizado um método hierárquico, no qual os grupos formam uma hierarquia em que dados dois grupos, quaisquer que eles sejam, ou são disjuntos ou um deles está contido no outro. Em particular, foi utilizado o método da ligação completa, de acordo com o qual a dissemelhança entre duas classes C_r e C_s é definida como sendo a maior dissemelhança entre um elemento de C_r e um elemento de C_s : $D_{C_r C_s} = \max\{d_{rs} : r \in C_r \text{ e } s \in C_s\}$.

O objetivo principal da classificação hierárquica consiste em construir uma árvore, cuja representação gráfica, a duas dimensões, se designa por dendrograma, diagrama de árvore, árvore hierárquica ou fenograma.

De forma a ilustrar a utilização desta técnica, considere-se o gráfico obtido para o exemplo abordado nas secções anteriores via Matriz Disjuntiva Completa e o qual abaixo se reproduz, de forma a facilitar a leitura.

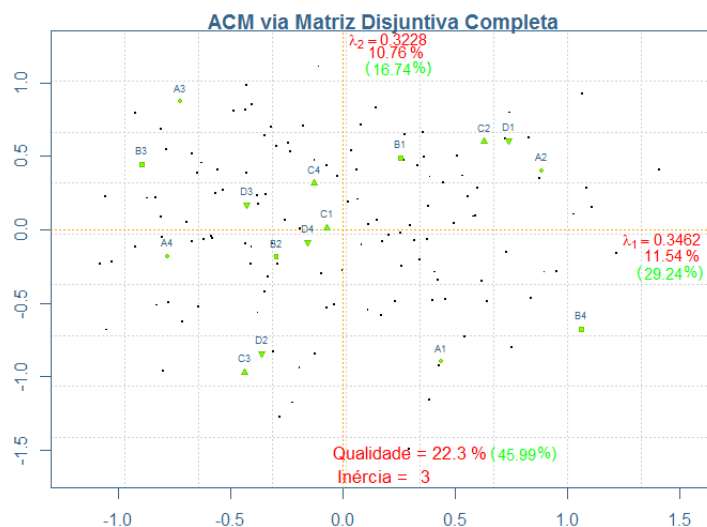


Figura 5.18: Réplica do gráfico apresentado na figura 5.13.

À partida, através da observação do gráfico 5.18, é possível identificar grupos de categorias fortemente relacionadas. São eles:

- Grupo I: Categorias C3 e D2;
- Grupo II: Categorias A2, C2 e D1;
- Grupo III: Categorias B2, C1, C4, D3 e D4.

No que respeita às restantes categorias não é trivial a definição dos grupos. Isto é, através da observação do gráfico apresentado na figura 5.18 não é possível identificar com exatidão a que grupo pertence cada uma das restantes categorias. Desta forma, a utilização da Análise Classificatória vem dar um contributo adicional à interpretação gráfica.

Após aplicar a Análise Classificatória, considerando a matriz de distâncias (dissemelhâncias) dada pelas coordenadas das categorias no mapa da ACM, através do método da ligação completa, obtém-se o dendrograma apresentado na figura 5.19.

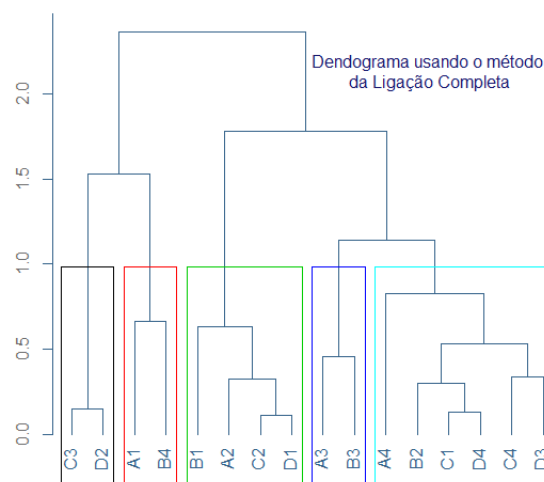


Figura 5.19: Dendrograma obtido através do Método da Ligação Completa aplicado à matriz de distâncias dada pelas coordenadas das categorias, no mapa apresentado na figura 5.18

Verifica-se que o maior salto é dado quando se agrupam os *clusters* (grupos) formados pelas categorias C3 e D2 (grupo 1) e A1 e B4 (grupo 2). Formam-se assim 5 grupos, formados pelas seguintes categorias:

- 1º grupo: C3 e D2;
- 2º grupo: A3 e B3;
- 3º grupo: B1, A2, C2 e D1;
- 4º grupo: A1 e B4;
- 5º grupo: A4, B2, C1, D4, C4 e D3

Repare-se que o 1º grupo corresponde precisamente ao grupo I que já havia sido identificado pela observação gráfica do mapa de ACM. Por outro lado, o 3º grupo corresponde ao grupo II, com a inclusão da categoria B1. Já o 5º grupo contém as categoria consideradas no grupo III, juntamente com a categoria A4.

Os restantes grupos (2º e 5º) são a partição das restantes categorias que suscitaram dúvidas aquando a interpretação exclusiva do mapa de ACM.

De forma a tornar-se claro os grupos formados, na figura 5.20 apresenta-se novamente o mapa resultante da ACM via Matriz Disjuntiva Completa, com os grupos devidamente identificados.

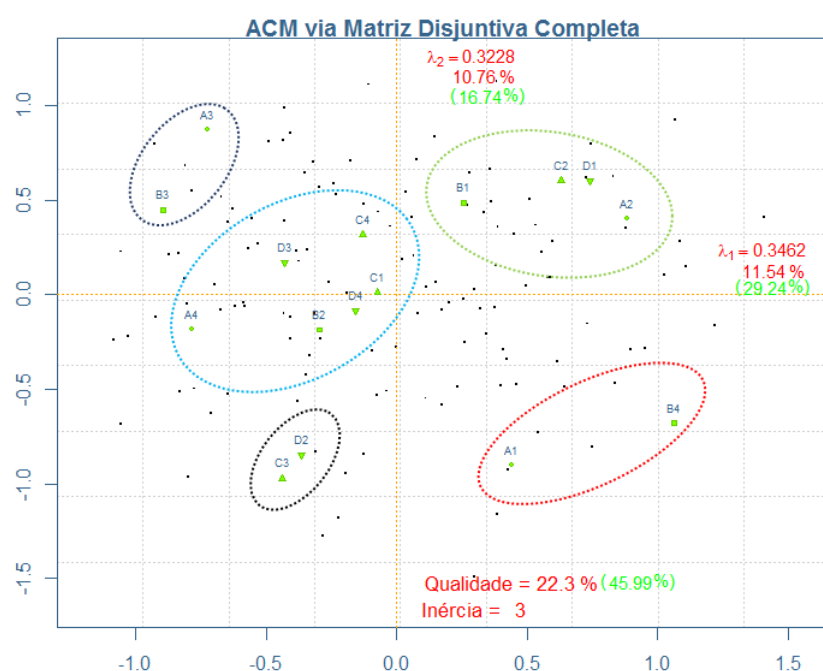


Figura 5.20: Mapa ACM com grupos (de categorias) identificados.

Com este simples exemplo, tornou-se clara a complementariedade entre estas duas metodologias de análise multivariada. Por meio da Análise Classificatória, foi possível agrupar as categorias que suscitaram dúvidas aquando a aplicação da ACM.

Para além disso, é importante referir que a construção dos grupos através da Análise Classificatória reflete, como era natural esperar, a posição das modalidades no gráfico da ACM.

No 1^o passo, forma-se o grupo com as categorias C2 e D1. Ora, se verificarmos a posição destas categorias no gráfico, verifica-se que são duas das categorias mais próximas entre si, a par de outras. De seguida, no segundo passo, é formado o grupo com as categorias C1 e D4, as quais se encontram bastante próximas. E, assim sucessivamente.

Repare-se ainda que, no grupo formado pelas categorias A4, B2, C1, C4, D3 e D4, a última categoria a entrar no *cluster* é, precisamente, o elemento mais afastado, A4, o qual entra no grupo no 11^o passo.

Abaixo apresenta-se o *output* que representa ordem sobre a qual as categorias são agrupadas.

[1,]	-10	-13	[6,]	-12	-15	[11,]	-4	8
[2,]	-9	-16	[7,]	-3	-7	[12,]	7	11
[3,]	-11	-14	[8,]	4	6	[13,]	3	10
[4,]	-6	2	[9,]	-5	5	[14,]	9	12
[5,]	-2	1	[10,]	-1	-8	[15,]	13	14

Capítulo 6

Resultados

6.1 Análise Exploratória

Como indicado no capítulo 3, realizou-se, inicialmente, uma análise descritiva univariada, a qual teve como objetivo caracterizar a amostra, através da análise das variáveis em estudo em função da sua escala de medida, com recurso a medidas de tendência central, dispersão e forma. Recorreu-se, também, à representação gráfica como ferramenta auxiliar.

Esta análise, bem como as representações gráficas, foram obtidas com recurso ao Microsoft Excel, ao software R, versão 3.0.1 (www.r-project.org) e ao software SPSS, versão 20.0.0.

A análise exploratória, ir-se-á centrar, primeiramente, nos potenciais fatores associados ao doente. Posteriormente, o foco de atenção serão os fatores relacionados com a situação clínica e, finalmente, os fatores referentes à relação com os serviços de saúde.

6.1.1 Fatores Relacionados com o Doente

1. Caracterização socio-demográfica

A amostra inclui 177 participantes com idades compreendidas entre os 23 e 85 anos. Os doentes apresentaram uma idade média de 48,4 com desvio-padrão de 10 anos, sendo que 50% tem menos de 47 anos. Mais de três quartos, 76,3% (N=135), é do sexo masculino. A grande maioria, 83,6% (N=148), nasceu em Portugal, é de nacionalidade portuguesa, 95,5% (N=169) e 91% (N=161) é de etnia branca de origem europeia. 15,9% (N=28) é

natural de países de língua oficial portuguesa (excluindo Portugal). 43,5% (N=77) dos participantes são solteiros, enquanto 37,3% (N=66) são casados (tabela 6.1).

Dos 177 participantes, 48,6% (N=86) foram classificados como não aderentes à terapêutica anti-retrovírica.

Tabela 6.1: Caracterização demográfica da amostra

Dados demográficos		
(N = 177)		
Idade [média±DP ¹ (anos)]		48,4±10,0
Sexo [%] (N)	Masculino	76,3 (135)
	Feminino	23,7 (42)
Naturalidade [% (N)]	Portugal	83,6 (148)
	Angola	4,0 (7)
	Moçambique	7,3 (13)
	Cabo-Verde	1,1 (2)
	Guiné-Bissau	2,3 (4)
	Outro ²	1,8 (3)
Nacionalidade [% (N)]	Portuguesa	95,5 (169)
	Angolana	1,1 (2)
	Outra ³	3,4 (6)
Etnia [% (N)]	Branco de origem europeia	91,0 (161)
	Branco de origem não-europeia	2,8 (5)
	Negro	3,4 (6)
	Mulato	2,8 (5)
Estado Civil [% (N)]	Solteiro	43,5 (77)
	Casado/União de facto	37,3 (66)
	Divorciado	14,1 (25)
	Viúvo	5,1 (9)

¹DP - Desvio-padrão.

²Brasil (0,6% (N=1)), São Tomé (0,6% (N=1)) e Inglaterra (0,6% (N=1)).

³guineense (0,6% (N=1)), brasileira (0,6% (N=1)), portuguesa/espanhola (0,6% (N=1)), portuguesa/angolana (0,6% (N=1)), inglesa (0,6% (N=1)) e portuguesa/cabo-verdiana (0,6% (N=1)).

No que toca à idade dos participantes, na figura 6.1, apresentam-se três representações gráficas usuais quando a variável em estudo é a idade.

De acordo com os gráficos apresentados, em particular, através do box-plot, é de salientar a simetria na zona central evidenciada pela estrutura da caixa. Todavia, observa-se uma assimetria à direita da distribuição das idades dos indivíduos em estudo [Coeficiente de assimetria¹ = 0,751].

¹O coeficiente de assimetria (ou *skewness*), como o próprio nome indica, avalia a (as)simetria da amostra.

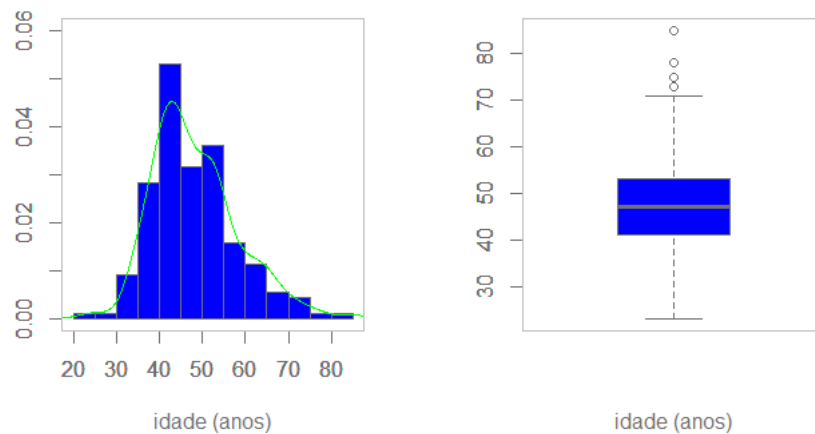


Figura 6.1: Densidade de Kernel e histograma (à esquerda) e Box-plot (à direita) das idades dos participantes.

Com o objetivo de verificar se a idade dos participantes segue a distribuição Normal, na figura 6.2, encontra-se representado o QQ-plot normal da distribuição das idades. Verifica-se que a reta não se ajusta convenientemente à nuvem de pontos, pelo que existem evidências para suspeitar de que a distribuição subjacente à idade dos participantes não segue uma normal.

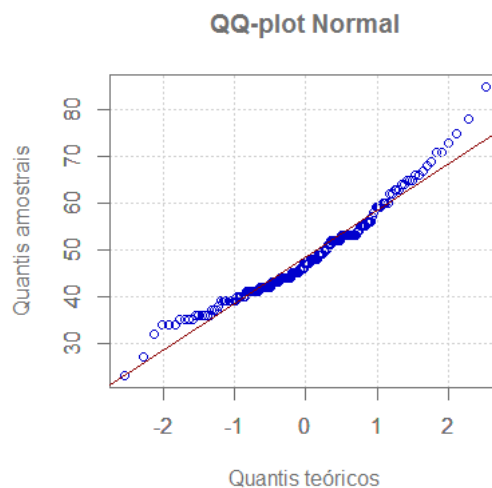


Figura 6.2: QQ-plot normal da distribuição das idades.

Realizando um teste estatístico, nomeadamente, o teste de Kolmogorov-Smirnov, com correção de Lilliefors, que testa as hipóteses

H_0 : A distribuição subjacente aos dados é normal (gaussiana)

vs

H_1 : A distribuição subjacente aos dados não é normal (gaussiana)

é rejeitada a hipótese de normalidade da distribuição das idades, uma vez que o valor observado da estatística de teste é $D=0,097$ e $\text{valor-p}=0,00$. Desta forma, o teste corrobora a análise gráfica acima efetuada.

Ainda de acordo com o Box-plot, observa-se a existência de possíveis outliers.

Com vista à aplicação da metodologia de ACM ao conjunto de dados, categorizou-se a variável idade da forma seguinte: $([18,35], [36,45], >45)$. Na figura 6.3 apresenta-se o respetivo gráfico de barras.

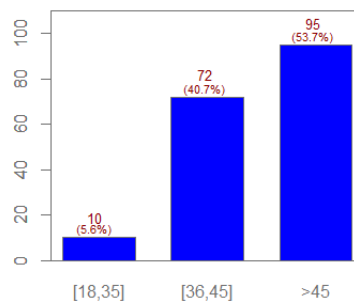


Figura 6.3: Gráfico de barras das categorias da variável idade.

Observa-se que existem 10 participantes entre os 18 e os 35 anos, 72 entre os 36 e os 45 e 95 com mais de 45 anos.

Com o objetivo de estudar a relação entre o sexo e a idade, na figura 6.4, apresentam-se os gráficos da densidade de Kernel ² das idades por sexo, bem como os box-plots das idades por sexo, sem e com intervalos de confiança a 95% para a mediana.

Com base nos gráficos apresentados, verifica-se um ligeiro aumento na idade dos indivíduos do género masculino, com uma acentuada assimetria à direita para a distribuição das idades dos homens. Apesar de não ser tão acentuada, existe, igualmente, uma ligeira

²O método de Kernel é um método não paramétrico para a estimação de curvas de densidade onde cada observação é ponderada pela distância em relação a um valor central, denominado, usualmente, por núcleo.

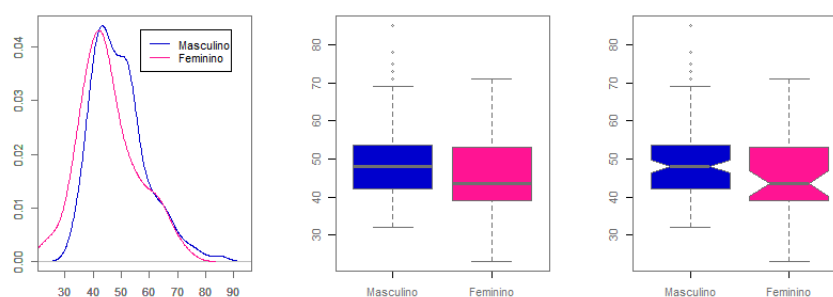


Figura 6.4: Densidade de Kernel das idades por género (à esquerda), boxplots das idades dos indivíduos por género (ao meio) e Boxplots das idades dos indivíduos por género com intervalos de confiança a 95% para a mediana (à direita).

assimetria à direita na distribuição das idades para o sexo feminino. De facto, o coeficiente de assimetria da distribuição das idades para o sexo masculino é 0,952, enquanto o coeficiente de assimetria da distribuição das idades para o sexo feminino é 0,47.

Para uma comparação mais eficiente, o R permite a representação do gráfico à direita, na qual, em cada caixa, é gerada uma chanfradura que representa um intervalo de confiança para as medianas. Quando as chanfraduras das caixas relativas aos diferentes grupos considerados não se sobrepõem, existem "fortes evidências" para afirmar que as medianas são diferentes (Chambers *et al.*, 1983). Neste caso, considerou-se um intervalo com 95% de confiança.

Com base no gráfico apresentado à direita (figura 6.4) parecem existir evidências para afirmar que existem diferenças acentuadas na distribuição das idades nos dois grupos, uma vez que as chanfraduras não se sobrepõem. Esta representação pode ser considerada um "teste" gráfico para as medianas.

No entanto, para averiguar se as distribuições das idades dos homens e das mulheres diferem entre si, realizou-se o teste de Mann-Whitney, o qual é utilizado com o objetivo de testar se as medianas dos dois grupos são iguais.

O valor-p do teste bilateral é de 0,024 (Estatística de teste $W=3491,5$), o qual, para um nível de significância de 5%, permite rejeitar a hipótese de igualdade de medianas. Conclui-se, assim, que há uma diferença significativa na distribuição da idade entre homens e mulheres, o que vai ao encontro da conclusão obtida através da análise gráfica.

Representando o QQ-plot normal da distribuição das idades para cada um dos géneros (figura 6.5), verifica-se que, no gráfico à direita, a nuvem de pontos encontra-se em torno da reta, o que não acontece com o QQ-plot Normal apresentado à esquerda, em particular,

para as idades mais baixas e mais altas, observa-se um afastamento dos pontos da reta.

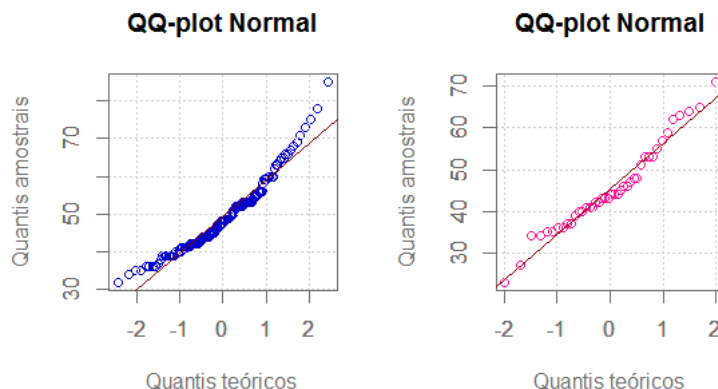


Figura 6.5: QQ-plot normal da distribuição das idades para o género masculino (à esquerda) e para o género feminino (à direita).

O teste de Kolmogorov-Smirnov, com correção de Lilliefors, corrobora a análise gráfica, uma vez que, para a distribuição das idades para o género masculino, é rejeitada a hipótese de normalidade ($D=0,102$, valor- $p=0,001$), enquanto, para a distribuição das idades para o género feminino, a hipótese de normalidade, ao nível de significância de 5%, não é rejeitada ($D=0,125$, valor- $p=0,096$).

No que respeita à habitação, 72,9% ($N=129$) dos participantes não mora sozinho. Destes, 16,4% ($N=29$) mora exclusivamente com o/a marido/esposa e 14,7% ($N=26$) mora com o/a marido/esposa e um(a) filho/filha. 90,6% ($N=116$) possui 4 co-habitantes ou menos. Apenas 18,6% ($N=33$) tem filhos menores a seu cargo (tabela 6.2).

Relativamente à escolaridade dos participantes, verifica-se que 14,7% ($N=26$) cumpriu apenas o 3º ciclo do ensino básico (5ºano do liceu (7-9 anos)), 20,3% ($N=36$) não avançou para além do secundário e 26,0% ($N=45$) graduou-se no ensino superior. Dos 177 participantes, 69,5% ($N=133$) cumpriram 9 ou mais anos de escolaridade com aproveitamento (tabela 6.3).

Ainda de acordo com a tabela 6.3, é possível verificar que, dos participantes empregados (56,5%), 7,6% ($N=8$) trabalha a partir de casa e 71,4% ($N=75$) tem um horário fixo. De notar que nenhum dos participantes é estudante.

2. Satisfação com o Suporte Social

Nas tabelas 6.4 e 6.5 apresentam-se os valores da consistência interna (Alfa de Cronbach) para a escala da Escala de Satisfação com o Suporte Social (ESSS), bem como de

Tabela 6.2: Caracterização dos co-habitantes dos participantes

Co-habitação		
(N = 177)		
Mora sozinho [%] (N)	Sim	27,1 (48)
	Não	72,9 (129)
Número de co-habitantes [mediana (AIQ)]		1 (2-3)
co-habitantes [%] (N)	Marido/Esposa	37,9 (67)
	Mãe	16,4 (29)
	Pai	5,6 (10)
	Sogra/Sogra	1,1 (2)
	Avô/Avó	1,7 (3)
	Filhos/Filhas	26,6 (47)
	Irmã/Irmão	4,5 (8)
	Outra pessoa da família	5,6 (10)
	Companheiro/a	5,6 (10)
	Enteado(s)	2,3 (4)
	Amigo(s)	3,4 (6)
	Outra pessoa ¹	4,8 (8)
Filhos menores a cargo [%] (N)	Sim	18,6 (33)
	Não	81,4 (144)

¹Outras pessoas que partilham a casa (0,6% (N=1)), comunidade terapêutica (0,6% (N=1)), nora (0,6% (N=1)), empregado (0,6% (N=1)), genro (0,6% (N=1)), companheiro da mãe (0,6% (N=1)), companheiro e amiga (0,6% (N=1)), albergue (0,6% (N=1)).

cada uma das suas subescalas, com o objetivo de avaliar a sua fiabilidade.

Segundo Marôco (2010), a classificação da consistência interna, de acordo com o valor do Alfa de Cronbach, pode ser obtida através da tabela 6.6.

Desta forma, com base na tabela 6.6, a ESSS apresenta uma consistência interna excelente ($\alpha = 0,903$). As dimensões satisfação com amigos, intimidade e satisfação com a família apresentam uma consistência interna boa ($\alpha = 0,897$, $\alpha = 0,716$ e $\alpha = 0,819$, respetivamente). Por outro lado, a subescala atividades sociais apresenta uma consistência interna aceitável ($\alpha = 0,696$).

No que diz respeito à correlação entre as diferentes subescalas, verifica-se, através das correlações de Spearman apresentadas na tabela 6.7, que todas as subescalas se encontram correlacionadas. As subescalas que apresentam menor correlação são satisfação com a família e satisfação com atividades sociais.

Como seria de esperar, a correlação entre as subescalas verifica-se no sentido positivo.

Tendo em conta que se trata de uma escala ordinal, foi ainda calculada a mediana

Tabela 6.3: Caracterização escolar e profissional dos participantes

Escolaridade e situação profissional		
(N = 177)		
Anos de escolaridade		9,69±4,79
[média±SD (anos)]		
	Nenhum (0 anos)	1,7 (3)
	1º ciclo do ensino básico	18,1 (32)
	2º Ciclo do ensino básico	10,7 (19)
	3º Ciclo do ensino básico	14,7 (26)
Nível de escolaridade [%]	Ensino secundário	20,3 (36)
	Ensino pós-secundário não superior	8,5 (15)
	Ensino superior universitário (bacharelato, licenciatura, mestrado e doutoramento)	26,0 (46)
	Desempregado	15,3 (27)
Situação laboral [%]	Reformado	23,8 (42)
	Empregado	56,5 (100)
	Doméstico (a)	3,4 (6)
	Outra ¹	1,7 (3)
(N = 105)		
Trabalho a partir de casa[%] ²	Sim	7,6 (8)
	Não	92,4 (97)
	Horário fixo	71,4 (75)
Horário de trabalho [%] ²	Horário por turnos	9,5 (17)
	Outro	19,0 (20)

¹Baixa médica (1,1% (N=2)) e Reformado+part-time (0,6% (N=1)).

² % de entre os que trabalham.

Tabela 6.4: Alfa de Cronbach para a escala ESSS no presente estudo

Escala	Alfa de Cronbach
ESSS - 15 itens	0,903

Tabela 6.5: Alfa de Cronbach para as subescalas da ESSS no presente estudo

Subescala	Alfa de Cronbach
Satisfação com amigos - 5 itens	0,897
Intimidade - 4 itens	0,716
Satisfação com a família - 3 itens	0,819
satisfação com atividades sociais - 3 itens	0,696

para cada uma das subescalas. A mediana obtida para as subescalas de satisfação com

Tabela 6.6: Classificação da consistência interna de acordo com o alfa de Cronbach

Alfa de Cronbach	Consistência interna
$\geq 0,9$	Excelente
$[0,7 - 0,9[$	Bom
$[0,6 - 0,7[$	Aceitável
$[0,5 - 0,6[$	Pobre
$< 0,5$	Inaceitável

Tabela 6.7: Correlação de Spearman entre as subescalas da ESSS

	Satisfação com amigos	Intimidade	Satisfação com a família	Satisfação com as atividades sociais
Satisfação com amigos				
Intimidade	$r=0,576^1$			
Satisfação com a família	$r=0,551^1$	$r=0,422^1$		
Satisfação com atividades sociais	$r=0,592^1$	$r=0,571^1$	$r=0,395^1$	

¹Correlação é significativa ao nível de significância de 1%.

amigos, intimidade, satisfação com a família e atividades sociais foi de 20, 14, 12 e 10, respetivamente.

Por outro lado, a pontuação máxima observada para cada uma das subescalas foi de 25, 20, 15 e 15, respetivamente. Isto significa que, para cada uma das dimensões, houve, pelo menos, um dos participantes que se mostrou totalmente satisfeito com o suporte social relativo à dimensão avaliada.

No que refere à pontuação (total) da ESSS, a mediana obtida foi de 46 e, à semelhança do que aconteceu com as subescalas, para a escala de ESSS, a pontuação máxima foi de 75, a qual corresponde ao *score* máximo possível, pelo que, houve, pelo menos, um participante que se revelou plenamente satisfeito com o suporte social.

Na identificação dos fatores associados à não-adesão, foi considerada apenas a pontuação final obtida para cada um dos participantes, isto é, a variável em estudo foi a satisfação com o suporte social, no seu sentido mais amplo.

3. Caracterização quanto à ansiedade, depressão e stress

No que respeita à escala de ansiedade, depressão e stress (EADS), nas figuras 6.8 e 6.9, é apresentado, respetivamente, o Alfa de Cronbach para a escala EADS e para cada uma das subescalas.

Tabela 6.8: Alfa de Cronbach para a escala EADS no presente estudo

Escala	Alfa de Cronbach
EADS - 21 itens	0,954

Tabela 6.9: Alfa de Cronbach para as subescalas do EADS no presente estudo

Subescala	Alfa de Cronbach
Ansiedade - 7 itens	0,869
Depressão - 7 itens	0,914
Stress - 7 itens	0,902

Em relação à correlação entre as diferentes subescalas, verifica-se, através das correlações de Spearman apresentadas na tabela 6.10, que todas as subescalas se encontram correlacionadas. Não existe uma correlação a destacar em relação às restantes.

Tabela 6.10: Correlação de Spearman entre as subescalas da EADS

	Ansiedade	Depressão	Stress
Ansiedade			
Depressão	$r=0,802^1$		
Stress	$r=0,786^1$	$r=0,810^1$	

¹Correlação é significativa ao nível de significância de 1%.

Tendo em conta que o intervalo de variação de cada uma das subescalas da EADS é o mesmo ([0,21]), faz sentido, neste caso, a representação dos box-plots paralelos, de forma a poder comparar a pontuação de cada uma das subescalas.

Com base nos box-plots apresentados na figura 6.6, facilmente se observa que a pontuação relativa ao stress é relativamente superior às restantes. Para além disso, através do gráfico à direita, verifica-se que a banda de confiança a 95% para a pontuação da subescala stress não se sobrepõe às restantes, pelo que, à partida, existe uma diferença significativa entre a mediana da pontuação da subescala stress e a mediana das restantes.

Porém, de forma a confirmar esta análise gráfica, realizou-se o teste de Kruskal-Wallis. Em primeiro lugar, procurou-se verificar se existe diferença significativa entre as pontuações das 3 subescalas, isto é, efetuou-se o seguinte teste:

H_0 : Não existe diferença significativa entre as pontuações das três subescalas

vs

H_1 : Existe diferença entre, pelo menos, duas subescalas

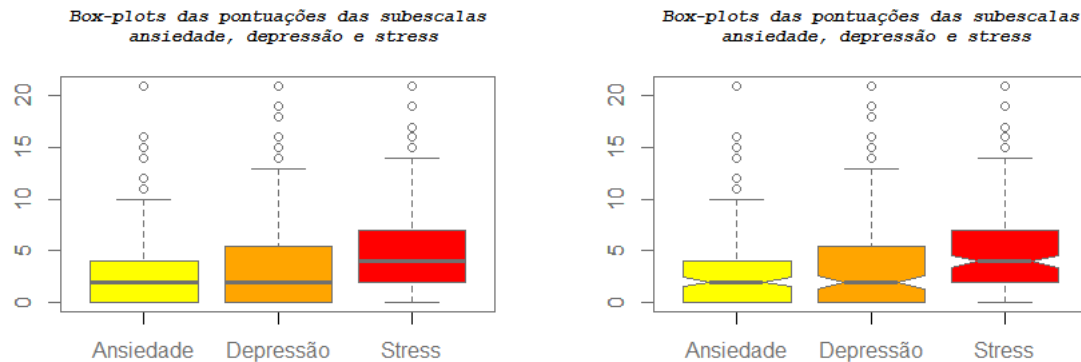


Figura 6.6: Box-plots das pontuações (à esquerda) e box-plot das pontuações com intervalos de confiança a 95% (à direita), das três escalas da EADS.

Obteve-se o seguinte *output*:

```
> kruskal.test(pontuação ~ fator)

Kruskal-Wallis rank sum test

data:  pontuação by fator
Kruskal-Wallis chi-squared = 37.5811, df = 2, p-value = 6.908e-09
```

Uma vez que o valor- $p=6,908e-09 < \alpha$ usual, tem-se que, para níveis de confiança usuais, em particular, $\alpha = 5\%$, rejeita-se a hipótese nula, isto é, rejeita-se a hipótese de não haver diferença significativa entre as pontuações das 3 subescalas. Não é, no entanto, possível, identificar qual(ais) o(s) grupo(s) que difere(m).

Com o objetivo de solucionar esta questão, o package *pgirmess* do R, possui uma função que permite verificar quais os grupos que diferem entre si. Desta forma, utilizando o package referido, obtém-se o seguinte *output*:

```
> kruskalmc(pontuação,fator)
Multiple comparison test after Kruskal-Wallis
p.value: 0.05
Comparisons
  obs.dif critical.dif difference
A-D 20.88424   39.02649      FALSE
A-S 94.13559   38.97117       TRUE
D-S 73.25135   39.02649       TRUE
```

Portanto, tal como já havia sido referido através da observação gráfica, efetivamente, existe uma diferença significativa entre a pontuação da subescala stress e cada uma das restantes subescalas.

Através dos box-plots verifica-se, igualmente, a existência de possíveis outliers nas três subescalas. Assim sendo, existem indivíduos que se destacam dos restantes por possuírem níveis elevados de depressão, ansiedade e stress.

A mediana das dimensões depressão, ansiedade e stress foi 2, 2, 4, respetivamente. Para além disso, tendo em conta que se trata de uma escala em que, quanto maior for a pontuação, mais negativo é o estado psico-patológico, faz sentido, determinar o mínimo obtido para cada subescala, o qual se verificou ser 0 para as três dimensões. Desta forma, existem indivíduos que não possuem qualquer sintoma de depressão, ansiedade e/ou stress.

Analisando, mais pormenorizadamente, a pontuação obtida para cada uma das subescalas, verifica-se que 28,2% (N=50) dos participantes não revelam qualquer indício de depressão (pontuação igual a 0), 27,7% (N=49) não apresentam sintomas de ansiedade e 11,3% (N=20) não apresenta sintomas de stress.

Finalmente, verifica-se ainda que, dos 177 participantes, 7,9% (N=14) não apresenta qualquer sintoma de depressão, ansiedade e stress (*score* igual a 0, para as três dimensões).

4. Outras características

Foram, igualmente, recolhidas variáveis referentes aos hábitos dos participantes, nomeadamente, regularidade na prática de exercício físico, consumo de álcool, hábitos tabágicos e outros consumos.

No que respeita aos hábitos de exercício físico, observa-se que, dos 177 participantes, 25,4% (N=45) praticam desporto diariamente e 23,2% (N=41) nunca pratica desporto. De entre os que praticam desporto, 63,8% (N=113) pratica mais do que 1, 2 vezes por semana (tabela 6.11).

Tabela 6.11: Caracterização dos hábitos de exercício físico dos participantes

Frequência de exercício físico		
Frequência de exercício físico [% (N)]	Diariamente	25,4 (45)
	mais de 4 vezes por semana	6,8 (12)
	3 a 4 vezes por semana	12,4 (22)
	1 a 2 vezes por semana	19,2 (34)
	2 a 3 vezes por mês	7,9 (14)
	1 vez por mês	5,1 (9)
	Nunca	23,2 (41)

Em relação à frequência no consumo de bebidas alcóolicas, verifica-se que 40,7% (N=72) dos participantes não consome álcool, sendo que, apenas 31,6% (N=56) consome mais do que 1 a 2 bebidas por semana (tabela 6.12).

Tabela 6.12: Caracterização dos hábitos de exercício físico dos participantes

Consumo de bebidas alcoólicas		
		(N = 177)
Frequência de consumo [% (N)]	Diariamente	12,4 (22)
	mais de 4 vezes por semana	1,1 (2)
	3 a 4 vezes por semana	7,9 (14)
	1 a 2 vezes por semana	10,2 (18)
	2 a 3 vezes por mês	14,7 (26)
	1 vez por mês	13,0 (23)
	Nunca	40,7 (72)
		(N = 105)
Bebidas por dia de consumo [% (N)]¹	1 a 2 bebidas por dia	78,1 (82)
	3 ou 4 bebidas por dia	11,4 (12)
	5 ou 6 bebidas por dia	6,7 (7)
	7 ou 8 bebidas por dia	2,9 (3)
	12 ou mais	1,0 (1)
Consumo de mais de 5 bebidas num período de 4 horas [%] (N)¹	Mais de 4 vezes por semana	1.9 (2)
	3 a 4 vezes por semana	1.0 (1)
	1 a 2 vezes por semana	7.6 (8)
	2 a 3 vezes por mês	9.5 (10)
	1 vez por mês	16.2 (17)
	Nunca	63.8 (67)

¹ % de entre os que consomem bebidas alcoólicas.

Para o presente estudo, os hábitos alcoólicos foram avaliados através de duas variáveis (Cohn *et al.*, 2011). A primeira variável, denominada de "Alcoolismo", reporta ao número médio de bebidas alcoólicas ingeridas por dia, nos trinta dias que antecederam a entrevista. Foram consideradas as seguintes categorias:

- **Sem hábitos alcoólicos:** Indivíduos que não ingeriram bebidas alcoólicas nos últimos 30 dias;
- **Hábitos ligeiros/moderados:** Indivíduos que ingeriram, em média, entre uma a quatro bebidas por dia, nos 30 dias que antecederam a entrevista;

- **Hábitos intensos:** Indivíduos que ingeriram, em média, mais do que quatro bebidas por dia, nos 30 dias anteriores à entrevista.

A segunda variável, denominada de "*Binge Drinking*"³ refere-se ao consumo excessivo de bebidas alcoólicas num curto espaço de tempo. Trata-se de uma variável dicotômica, onde os indivíduos são classificados como *binge drinkers*, caso tenham reportado um consumo de cinco ou mais bebidas num espaço de quatro horas. Se referiram que, nos últimos 30 dias, nunca consumiram cinco ou mais bebidas alcoólicas num espaço de quatro horas foram considerados como não sendo *binge drinkers*.

De acordo com estas duas variáveis, verifica-se que 36,7% (N=65) dos participantes apresentam hábitos alcoólicos intensos, isto é, consumiram, em média, mais do que quatro bebidas por dia, nos 30 dias que antecederam a entrevista (tabela 6.13). Por outro lado, dos 177 participantes, 21,5% (N=38) foram considerados *binge drinkers* (tabela 6.14).

Tabela 6.13: Classificação dos hábitos alcoólicos de acordo com o número médio de bebidas ingeridas por dia, nos últimos 30 dias

Alcoolismo			
			(N = 177)
Hábitos de consumo [%] (N)	Sem hábitos		40,7 (72)
	Hábitos ligeiros/moderados		22,0 (39)
	Hábitos intensos		37,3 (66)

Tabela 6.14: Classificação de *binge drinker*

<i>Binge Drinking</i>			
			(N = 177)
<i>Binge drinker</i> [%] (N)	Sim	21,5 (38)	
	Não	78,5 (139)	

Dos indivíduos que possuíam hábitos alcoólicos ligeiros/moderados, 21,6% (N=9) foram considerados *binge drinkers* e, de entre os que possuíam hábitos intensos, 42,5% (N=29) foram considerados *binge drinkers* (figura 6.7).

Como seria de esperar, as duas variáveis encontram-se fortemente relacionadas ($X^2=36,44$, valor-p= $1,22 \cdot 10^{-8}$).

³Esta é a expressão utilizada para descrever o consumo excessivo de álcool que corresponde à ingestão de cinco ou mais bebidas alcoólicas num único dia ou momento, habitualmente, ao fim de semana.

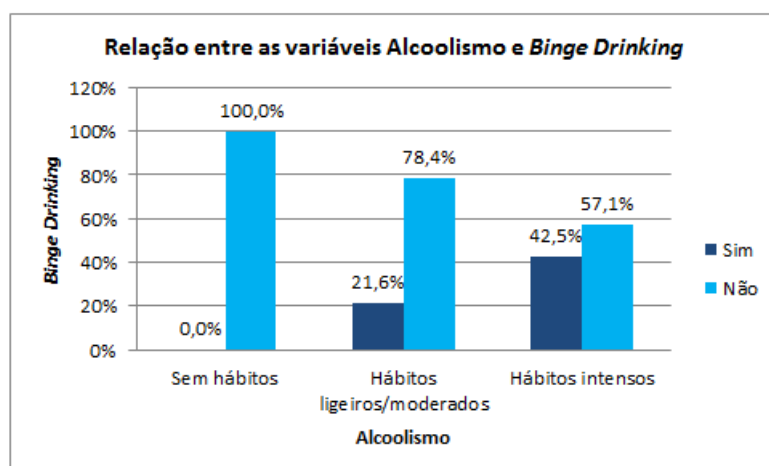


Figura 6.7: Relação entre as variáveis Alcoolismo e Binge Drinking

Em relação aos hábitos tabágicos verifica-se que 49,1% (N=87) dos participantes fuma e os restantes 50,9% (N=90) não fuma ou deixou de fumar (tabela 6.15).

Ainda de acordo com a tabela 6.15, verifica-se que, no que respeita ao consumo de drogas (leves e pesadas), verifica-se que, nos últimos 6 meses, 14,7% (N=26) dos participantes consumiu marijuana, 5,1% (N=9) consumiram cocaína e 1,1% (N=2) consumiram heroína.

Tabela 6.15: Caracterização do consumo de cigarros e drogas

Consumo de tabaco e drogas		
(N = 177)		
Consumo de tabaco [%] (N)	Nunca fumou	26,6 (47)
	Deixou de fumar	24,3 (43)
	Fuma ocasionalmente	9,6 (17)
	Fuma	39,5 (70)
Cigarros por semana (de entre quem fuma ocasionalmente)		17,88±10,05
[média±SD]		
Cigarros por dia (de entre quem fuma)		10,06±10,04
[média±SD]		
Droga [%] (N)	Marijuana	14,7 (26)
	Cocaína	5,1 (9)
	Heroína	1,1 (2)

Para além destas características foi ainda questionado aos indivíduos se já haviam estado em tratamento com metadona e, em caso afirmativo, se, no momento da entrevista, estavam em tratamento sob metadona. Na tabela 6.16 pode-se verificar que 10,7% (N=19) tinham estado anteriormente sob tratamento com metadona e destes, apenas 4,5% (N=8) estavam sob tratamento com metadona.

Tabela 6.16: Tratamento sob metadona

Tratamento sob metadona	
	(N = 177)
Anteriormente [% (N)]	10,7 (19)
	(N = 19)
Atualmente [% (N)]	4,5 (8)

No que respeita ao meio de transmissão do vírus, 83,6% (N=159) dos participantes indica como possível causa de transmissão a via sexual. Destes, 54,8% (N=97) reportam-se a uma relação sexual heterossexual e os remanescentes 28,8% (N=51) a uma relação sexual com homem (tabela 6.17).

Tabela 6.17: Modo de transmissão do vírus

Transmissão do vírus		
		(N = 177)
	Heterossexuais	45,2 (80)
	Homens que tiveram sexo com homens/Bissexuais	28,2 (50)
Possível meio de transmissão [%] (N)	Utilizadores de drogas injetáveis	15,3 (27)
	Heterossexuais/ Utilizadores de drogas injetáveis	5,6 (10)
	Outra via de transmissão	5,6 (10)

6.1.2 Fatores Relacionados com a Situação Clínica

Foram recolhidas algumas variáveis, tais como a adesão estimada pelo médico, a presença/ausência de co-infecções e co-morbilidades, o tempo até iniciar a TAR, o tempo sob TAR e o tempo de seguimento, entre outras.

Dos 177 participantes, 95,5% (N=169) não se encontram com classificação SIDA, 75,1% (N=133) não apresentam co-infecções e 68,9% (N=122) não registam a presença de comorbilidades.

Ainda de acordo com o indicado pelo médico assistente, 92,1% (N=163) dos doentes aderem à terapêutica anti-retrovírica (adesão igual ou superior a 95%). Dos participantes considerados não aderentes à terapêutica, 85,7% (N=12) foram considerados aderentes moderados (adesão superior ou igual a 50% e inferior a 95%) e 14,3% (N=2) são fracamente aderentes (adesão inferior a 50%).

Relativamente ao tempo (em meses) até iniciar TAR, sob TAR e em seguimento, verifica-se que o tempo médio sob terapêutica é, aproximadamente, de 10 anos e 2 meses (tabela 6.18).

Tabela 6.18: Tempo até iniciar TAR, sob TAR e de seguimento (em meses)

Variável	[média±DP]	[mediana(Q1-Q3) ¹]	N
Tempo até início de TAR	19,42±32,86	3,71 (0,53-25)	173
Tempo sob TAR	122,14±63,95	138,30 (68,11-165,15)	174
Tempo de seguimento	141,84±65,56	145,12 (94,98-189,54)	175

¹Q1-1º quartil, Q3-3º quartil.

Foi ainda solicitado aos médicos que indicassem a classificação CDC do doente aquando a primeira consulta, no momento do início da TAR e aquando a entrevista.

Este sistema, globalmente aceite para a classificação da infeção por VIH, inicialmente publicado pelo U.S. Centers of Disease Control and Prevention em 1986, baseia-se em critérios específicos que estão associados à infeção.

A atual classificação do CDC classifica os indivíduos com base nas condições clínicas e na contagem de linfócitos T CD4. Existem três categorias clínicas:

- **A:** doentes assintomáticos;
- **B:** doentes sintomáticos;
- **C:** doentes com classificação SIDA;

e três categorias para os linfócitos:

- **1:** > 500 células/ μ l;
- **2:** 200 – 499 células/ μ l;

- **3:** < 200 células/ μl ;

Uma vez classificados numa determinada categoria, os indivíduos não podem, no futuro, ser classificados numa categoria inferior (Castro *et al.*, 1992). Por exemplo, um indivíduo que, num determinado momento, é classificado como B1 (sintomático e contagem de linfócitos entre 200 e 499 células/ μl) não pode, futuramente, ser classificado como A1 (assintomático e contagem de células T CD4 entre 200 e 499 células/ μl).

A tabela 6.19 apresenta a proporção de indivíduos em cada um dos estados CDC, para os três momentos considerados.

Tabela 6.19: Classificação CDC dos indivíduos nos três momentos avaliados

Classificação CDC					
1ª consulta		Início da TAR		Data da entrevista	
A1 [%](N)	24,7 (43)	A1 [%](N)	11,5 (20)	A1 [%](N)	8,0 (14)
A2 [%](N)	27,0 (47)	A2 [%](N)	31,6 (55)	A2 [%](N)	31,6 (55)
A3 [%](N)	13,2 (23)	A3 [%](N)	16,7 (29)	A3 [%](N)	16,7 (29)
B1 [%](N)	4,0 (7)	B1 [%](N)	3,4 (6)	B1 [%](N)	2,9 (5)
B2 [%](N)	5,7 (10)	B2 [%](N)	7,5 (13)	B2 [%](N)	7,5 (13)
B3 [%](N)	9,2 (16)	B3 [%](N)	11,5 (20)	B3 [%](N)	11,5 (20)
C1 [%](N)	0,6 (1)	C1 [%](N)	0,6 (1)	C1 [%](N)	0,6 (1)
C2 [%](N)	1,7 (3)	C2 [%](N)	1,7 (3)	C2 [%](N)	2,3 (4)
C3 [%](N)	13,8 (24)	C3 [%](N)	17,2 (30)	C3 [%](N)	20,7 (36)
Valores omissos [%](N)	1,7% (3)	Valores omissos [%](N)	—	Valores omissos [%](N)	—

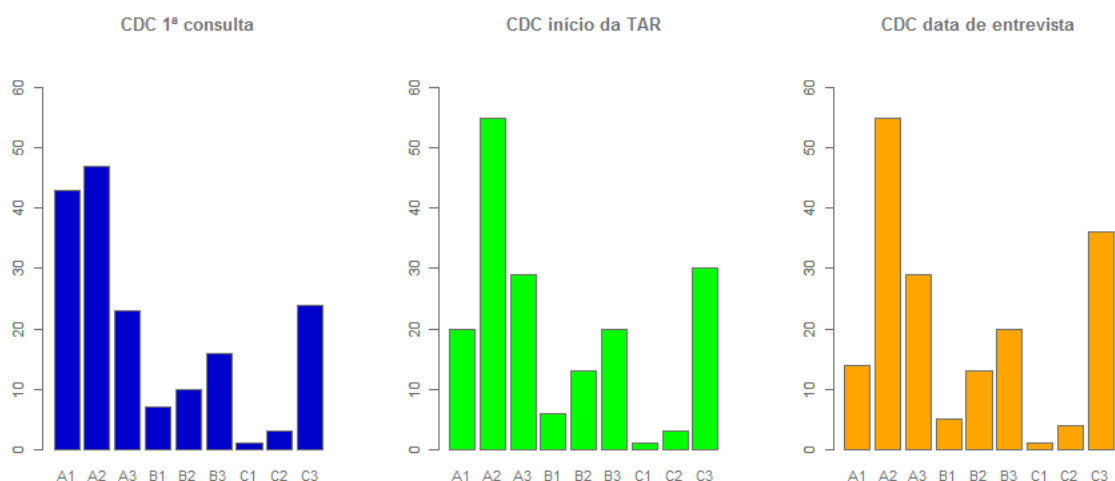


Figura 6.8: Gráfico de barras relativos à classificação CDC nos três momentos avaliados.

6.1.3 Fatores Relacionados com as Crenças sobre a Doença e os Fármacos

1. Revised Illness Perception Questionnaire

Em relação ao instrumento *Revised Illness Perception Questionnaire* - (*IPQ-R*), na figura 6.20 é apresentado o Alfa de Cronbach obtido com a amostra em estudo para cada uma das subescalas que compõem a referida escala.

Tabela 6.20: Alfa de Cronbach para as subescalas do IPQ-R no presente estudo

Subescala	Alfa de Cronbach
Identidade - 14 itens	0,870
Duração aguda/crónica - 6 itens	0,707
Duração cíclica - 4 itens	0,869
Consequências - 6 itens	0,764
Controlo pessoal - 6 itens	0,612
Controlo de tratamento - 5 itens	0,539
Coerência da doença - 5 itens	0,725
Representação emocional - 6 itens	0,838

As subescalas identidade, duração aguda/crónica, duração cíclica, consequências, coerência da doença e representação emocional apresentam uma consistência interna boa ($\alpha=0,87$, $\alpha=0,707$, $\alpha=0,869$, $\alpha=0,764$, $\alpha=0,725$ e $\alpha=0,838$, respetivamente). Por sua vez, as subescalas controlo pessoal e controlo de tratamento apresentam uma consistência interna aceitável ($\alpha=0,612$) e pobre ($\alpha=0,539$), respetivamente.

Com o objetivo de simplificar a escrita, considere-se, de seguida, os seguintes acrónimos: I - Identidade, DAC - Duração Aguda/Crónica, DC - Duração Cíclica, C - Consequências, CP - Controlo Pessoal, CT - Controlo de Tratamento, CD - Coerência da Doença e RE - Representação Emocional.

Na figura 6.21 apresentam-se as correlações de Spearman entre as subescalas do instrumento *IPQ-R*, das quais se destacam as seguintes:

- **$r=0,505$** entre as subescalas controlo pessoal e controlo de tratamento:

Recorde-se que, no que se refere à subescala de controlo pessoal, quanto maior for a pontuação obtida, maior é a crença do doente sobre o papel fulcral do próprio na minimização do impacto da doença. Isto é, o doente acredita que tem um papel

Tabela 6.21: Correlação de Spearman entre as subescalas do IPQ-R

	I	DAC	DC	C	CP	CT	CD	RE
I								
DAC	r=0,174 ²							
DC	r=0,344 ³	r=-0,018 ¹						
C	r=0,356 ³	r=0,410 ³	r=0,342 ³					
CP	r=0,069 ¹	r=0,335 ³	r=-0,256 ³	r=0,219 ³				
CT	r=-0,162 ³	r=0,046 ¹	r=-0,207 ²	r=0,040 ¹	r=0,505 ³			
CD	r=0,212 ³	r=-0,162 ¹	r=0,406 ³	r=0,157 ¹	r=-0,299 ³	r=-0,401 ³		
RE	r=0,423 ³	r=0,138 ¹	r=0,449 ³	r=0,506 ³	r=-0,029 ¹	r=-0,092 ¹	r=0,440 ³	

¹Correlação não significativa.²Correlação é significativa ao nível de significância de 5%.³Correlação é significativa ao nível de significância de 1%.

ativo no controlo da doença. Já no que diz respeito à subescala de controlo de tratamento, quanto maior for a pontuação observada, maior é a perceção do doente sobre a necessidade e eficácia do tratamento. O doente acredita que o tratamento constitui um papel fundamental na cura/controlo da doença.

Desta forma, a correlação positiva tem sentido, uma vez que, ao acreditar que o tratamento desempenha um papel fundamental na cura/controlo da doença, o indivíduo compreende que para que o tratamento surta efeito, deve cumprir os regimes terapêuticos tal como foram prescritos, de forma a obter a maior eficácia possível sob tratamento. Assim, o participante acredita que as suas ações, nomeadamente, não falhar a toma da medicação, têm implicações diretas com a cura/controlo da doença. Por contraposição, quando o doente não acredita no tratamento, crê que não há nada a fazer para controlar a doença, pelo que se sente impotente para combater a doença;

- **r=0,506** entre as subescalas representação emocional e consequências:

A subescala representação emocional diz respeito ao estado emocional e psicológico dos indivíduos face à doença. Pontuações mais elevadas correspondem a estados mais negativos.

Por sua vez, a subescala de consequências está relacionada com a perceção dos indivíduos sobre a gravidade da doença. À semelhança das restantes subescalas, pontuações mais elevadas correspondem a perceções mais negativas.

Assim sendo, a correlação positiva entre estas duas subescalas é facilmente interpretável, uma vez que, quanto maior a crença sobre a gravidade da doença, maior é a

depressão e preocupação sobre a doença.

- $r = -0,401$ entre as subescalas coerência da doença e controlo de tratamento:

Em relação à subescala coerência da doença, a mesma reflete o (des)conhecimento acerca da doença. Quanto maior a pontuação, maior é a falta de informação sobre a doença.

Desta forma, a correlação negativa facilmente se explica, uma vez que quanto maior o desconhecimento sobre a doença, a gravidade da doença é relevada.

A mediana obtida para cada uma das subescalas foi 3 (identidade), 24 (duração aguda/crónica), 10 (duração cíclica), 21 (consequências), 23 (controlo pessoal), 19 (controlo de tratamento), 11 (coerência da doença) e 18 (representação emocional).

Por outro lado, a pontuação mínima obtida para cada uma das dimensões foi 0 (identidade), 12 (duração aguda/crónica), 4 (duração cíclica), 10 (consequências), 13 (controlo pessoal), 11 (controlo de tratamento), 5 (coerência da doença) e 6 (representação emocional).

Recordando o intervalo de variação de cada uma das dimensões, mencionado na secção 3.1.3, observa-se que, existe, pelo menos, um indivíduo que obteve a pontuação mínima na subescala identidade. O mesmo se passou para as subescalas duração cíclica, coerência da doença e representação emocional. Desta forma, há pelo menos um participante que não associa qualquer sintoma à sua doença, bem como existe, pelo menos um participante que atribui carácter periódico algum à doença. Verifica-se, igualmente, que existe, no mínimo, um participante, que entende possuir uma informação vasta sobre a doença e, por fim, há, pelo menos, um indivíduo, que não apresenta qualquer representação emocional negativa acerca da doença.

Observando em detalhe o *score* obtido para cada um dos 177 indivíduos, constata-se que 29,4% (N=52) dos participantes não relacionam qualquer um dos catorze sintomas à doença, 6,2% (N=11) dos participantes não atribuem carácter cíclico algum à doença, 4,0% (N=7) dos indivíduos, aparentemente, compreendem totalmente a doença e apenas 0,6% (N=1) dos doentes não apresenta representação emocional negativa sobre a doença.

Na figura 6.9 representam-se os box-plots das pontuações obtidas para cada uma das subescalas, nos quais se pode observar a dispersão da pontuação. Por exemplo, o mínimo da subescala identidade é 0 e o 3º quartil é 6, pelo que, 75% dos participantes obtiveram uma pontuação entre 0 e 6, e destes, 50%, ficaram abaixo do valor 3 (mediana).

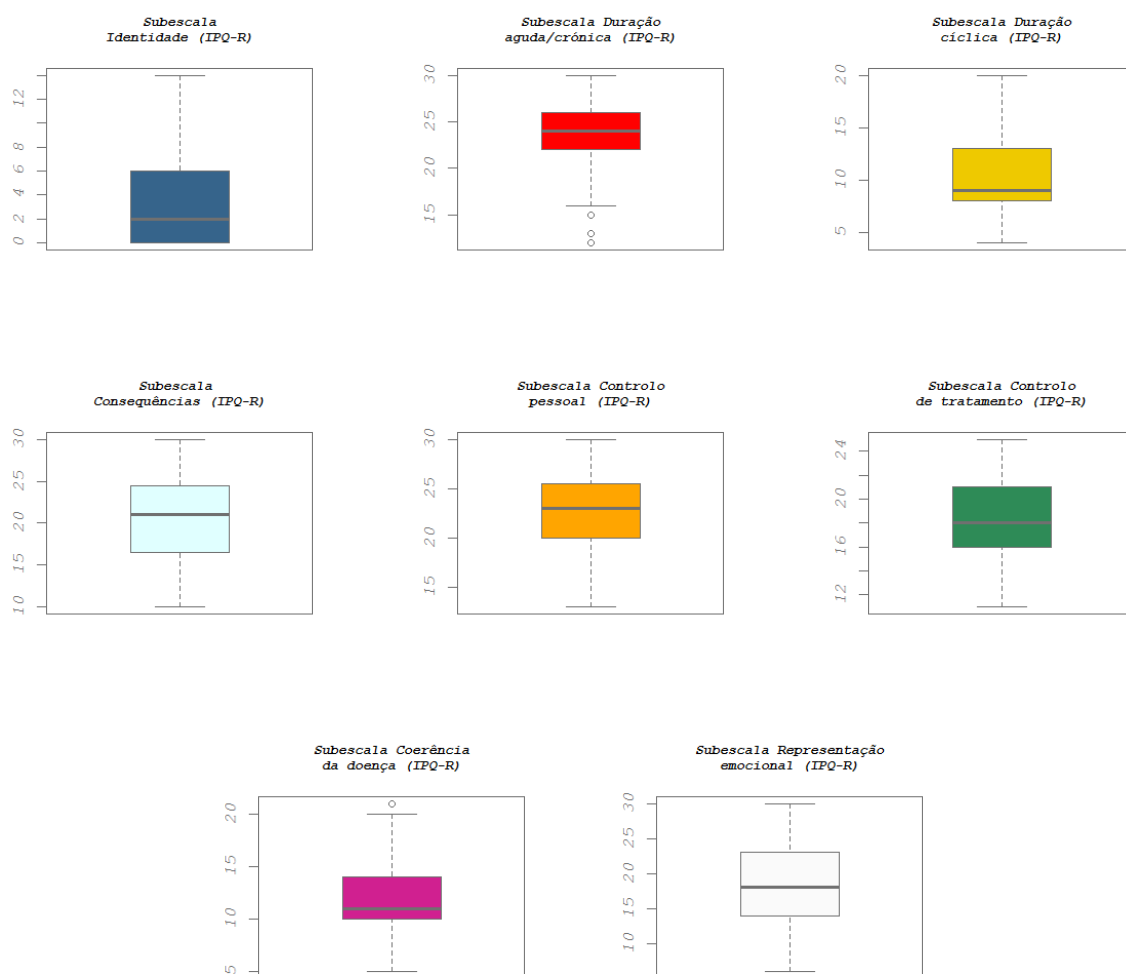


Figura 6.9: Box-plots das pontuações das subescalas do *IPQ-R*

Observa-se ainda a existência de possível simetria na pontuação das dimensões de consequências, controle pessoal, controle de tratamento e representação emocional.

2. Questionário de Crenças Acerca da Medicação

Por último, no que concerne ao instrumento de crenças acerca da medicação, nas figuras 6.22 e 6.23 apresenta-se o Alfa de Cronbach para a escala e para cada uma respectivas subescalas.

Tabela 6.22: Alfa de Cronbach para a escala BMQ no presente estudo

Escala	Alfa de Cronbach
BMQ - 19 itens	0,666

Tabela 6.23: Alfa de Cronbach para as subescalas do BMQ no presente estudo

Subescala	Alfa de Cronbach
Preocupações - 11 itens	0,768
Necessidades - 8 itens	0,811

De acordo com as tabelas 6.22 e 6.23, verifica-se que a consistência interna das subescalas do instrumento BMQ é boa ($\alpha = 0,768$ e $\alpha = 0,811$). Por outro lado, a consistência interna para a escala revela-se aceitável ($\alpha = 0,666$).

Tabela 6.24: Correlação de Spearman entre as subescalas do BMQ

	Preocupações	Necessidades
Preocupações		
Necessidades	$r=-0,205^1$	

¹Correlação é significativa ao nível de significância de 1%.

A subescala de necessidades, tal como já fora mencionado no capítulo 3, avalia a percepção do indivíduo sobre a necessidade da medicação no controlo da doença. Por outro lado, a subescala preocupações mede o nível de preocupação do indivíduo acerca da medicação, como, por exemplo, a manifestação de efeitos secundários, sabor da medicação, entre outros.

Recorde-se que, para a subescala de necessidades, quanto maior a pontuação obtida, maior é a crença do indivíduo acerca da necessidade da medicação. No que respeita à subescala de preocupações, pontuações mais elevadas refletem maiores preocupações sobre a medicação.

Desta forma, a correlação de $r=-0,205$ vai ao encontro do esperado, uma vez que, acreditando que a medicação é absolutamente necessária para a contenção da doença, o indivíduo, à partida, não manifesta preocupações sobre a mesma.

A mediana para cada uma das subescalas foi 25 (preocupações) e 32 (necessidades). Por outro lado, a pontuação máxima obtida para cada uma das dimensões foi, respetivamente, 38 e 40. Desta forma, tendo em conta que pontuação da subescala necessidades varia no intervalo [8,40], verifica-se que existem indivíduos que entendem a medicação como absolutamente necessária para a cura/contenção da doença. Mais concretamente, 4,5% (N=8) participantes, acreditam que a medicação anti-retrovírica é fundamental no seu bem-estar.

Tendo em conta que a pontuação das subescalas preocupações e necessidades varia entre [11,55] e [8,40], respetivamente, tem-se que o ponto médio dos respetivos intervalos

é 33 e 24. Assim sendo, facilmente se verifica que, em média, os participantes não revelam preocupações acentuadas sobre a medicação e revelam estar cientes acerca da necessidade da medicação.

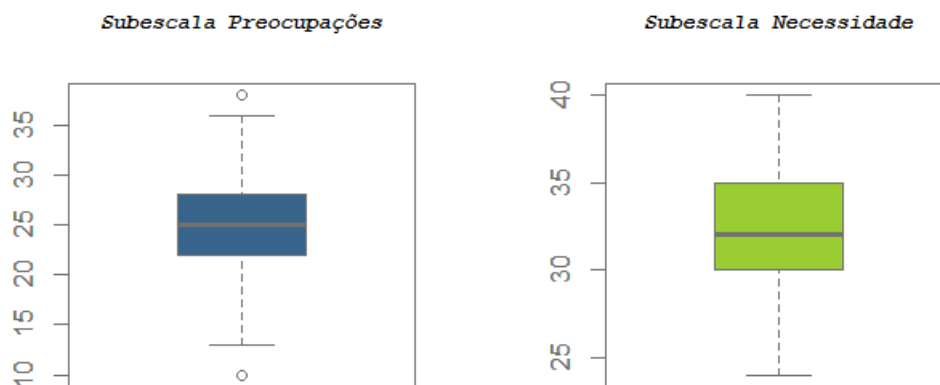


Figura 6.10: Box-plots das subescalas do instrumento BMQ.

Na figura 6.10 apresentam-se os box-plots referentes às pontuações obtidas para cada uma das subescalas, nos quais se pode observar a dispersão das pontuações. No que respeita à subescala preocupações, verifica-se a existência de possíveis outliers inferiores e superiores.

6.2 Análise Bivariada

Como foi referido no capítulo 3, com vista a avaliar a associação entre as variáveis preditoras e a não-adesão foi, inicialmente, aplicado o teste do qui-quadrado (ou teste exato de Fisher) para verificar a existência de associação entre as variáveis categóricas e a não-adesão e o teste de Mann-Whitney para comparar as medianas das variáveis contínuas entre os grupos dos indivíduos aderentes e dos indivíduos não aderentes.

Na tabela 6.25 apresentam-se as variáveis que se revelaram associadas à não-adesão (valor- $p \leq 0,05$).

Tabela 6.25: Variáveis que se verificaram relacionadas com a não-adesão, aquando aplicação do teste do qui-quadrado (ou teste exato de Fisher) ou teste de Mann-Whitney

Variável	Valor-p
Morar sozinho	0,021
<i>Binge drinking</i>	0,000
Ansiedade (EADS)	0,006
Recomendação sobre TAR	0,010
Recomendação sobre adesão	0,028
Identidade (<i>IPQ-R</i>)	0,017
Controlo de tratamento (<i>IPQ-R</i>)	0,047
Representação emocional (<i>IPQ-R</i>)	0,037
Preocupações (BMQ)	0,006

6.2.1 Fatores Relacionados com o Doente

Através da tabela 6.25 verifica-se que, de entre as variáveis relacionadas com o doente, observou-se que existe associação entre a variável resposta e o facto de morar ou não sozinho (valor-p=0,0219), o facto de ser ou não *binge drinker* (valor-p=0,000) e o estado de ansiedade (valor-p=0,006).

É, no entanto, de referir que, para as variáveis alcoolismo e estado de depressão o valor-p obtido foi de 0,053 e 0,057, respetivamente. Trata-se, portanto, de *borderlines*, isto é, valores-p próximos do nível de significância considerado, 0,05.

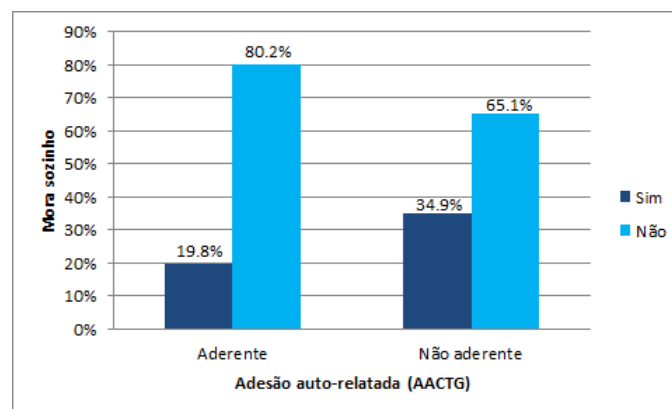


Figura 6.11: Comparação entre os grupos dos participantes aderentes e não aderentes à terapêutica no que se refere ao facto de morarem ou não sozinhos

Através do gráfico apresentado na figura 6.11, observa-se que, dos indivíduos aderentes à terapêutica, 19,8% (N=18) vivem sozinhos, enquanto que, de entre o indivíduos não aderentes à TAR, 34,9% (N=30) vivem sozinhos. Verifica-se, portanto, que a proporção de

indivíduos que vivem sozinhos é superior no grupo dos doentes não aderentes à terapêutica anti-retrovírica.

Com base no gráfico apresentado na figura 6.12, verifica-se que, de entre os participantes classificados como aderentes, (apenas) 9,9% (N=9) reportam comportamentos de consumo excessivo de álcool num curto espaço de tempo. Quando considerados os indivíduos não aderentes à terapêutica, observa-se que essa percentagem aumenta para os 33,7% (N=29).

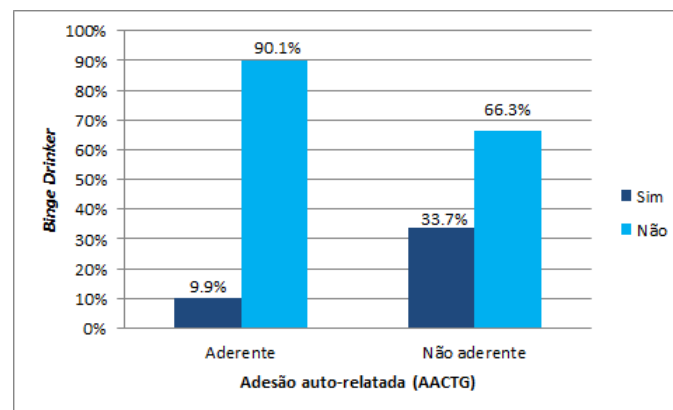


Figura 6.12: Comparação entre os grupos dos participantes aderentes e não aderentes à terapêutica no que se refere ao facto de serem ou não *binge drinkers*

Observando a figura 6.13, verifica-se que pontuações mais elevadas de ansiedade mostraram-se associadas à não adesão.

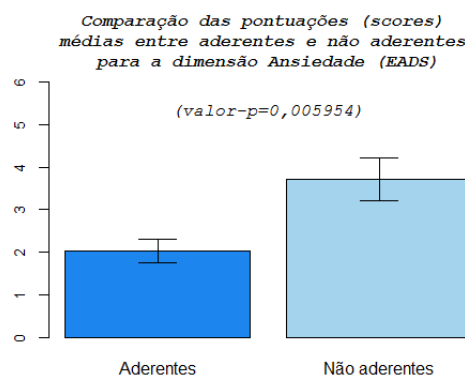


Figura 6.13: Comparação da pontuação (score) média entre aderentes e não aderentes para a dimensão ansiedade, avaliada pela EADS

6.2.2 Fatores Relacionados com a Situação Clínica

De entre as variáveis referentes à situação clínica do doente, observou-se associação entre a recomendação sobre TAR efetuada pelo médico assistente e a não-adesão (valor- $p=0,010$) e entre a recomendação de adesão efetuada pelo médico assistente (valor- $p=0,028$).

De acordo com o gráfico apresentado na figura 6.14, verifica-se que a proporção de indivíduos não aderentes foi maior entre os indivíduos aos quais foi efetuada recomendação sobre a TAR pelo médico assistente. Para além disso, com base no mesmo gráfico, observa-se que, para o grupo dos indivíduos não aderentes, a percentagem aos quais foi feita recomendação sobre a TAR é o dobro da percentagem aos quais não foi efetuada recomendação alguma. Por sua vez, para o grupo dos aderentes, esta percentagem é muito semelhante.

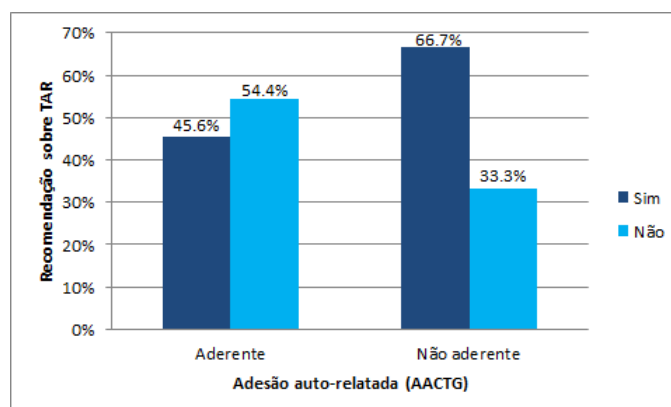


Figura 6.14: Comparação entre a percentagem doentes aderentes e não aderentes à terapêutica no que se refere à recomendação sobre TAR feita pelo médico

Em relação à recomendação de adesão efetuada pelo médico, através do gráfico da figura 6.15, facilmente se observa que, para o grupo dos doentes aderentes, a proporção de indivíduos aos quais foi recomendada a adesão à TAR é (ligeiramente) inferior à proporção de indivíduos aos quais não foi efetuada recomendação. No que respeita aos indivíduos não aderentes à terapêutica, observa-se uma tendência contrária, isto é, a proporção de indivíduos aos quais foi recomendada adesão à TAR é substancialmente superior à proporção de indivíduos aos quais não foi efetuada recomendação.

De referir que, teoricamente, já se esperava este comportamento, uma vez que é natural que, perante um indivíduo não aderente, o médico sinta necessidade de recomendar a adesão à terapêutica. Não é de mais relembrar que, o sucesso da terapêutica anti-retrovírica depende fundamentalmente de um elevado nível de adesão, pelo que, os profissionais de

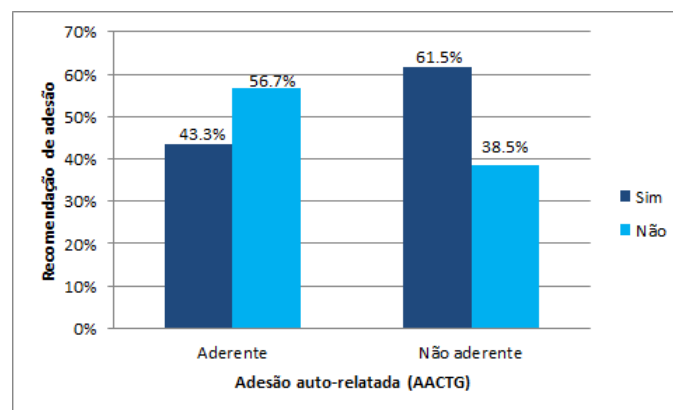


Figura 6.15: Comparação entre a percentagem de doentes aderentes e não aderentes à terapêutica no que se refere à recomendação de adesão à TAR feita pelo médico

saúde evidam esforços para que os doentes cumpram o regime terapêutico.

6.2.3 Fatores Relacionados com as Crenças sobre a Doença e os Fármacos

De entre os fatores relacionados com a perceção da doença e dos medicamentos, verificou-se, através do teste de Mann-Whitney, que a pontuação (*score*) média das dimensões identidade (valor- $p=0,017$), controlo de tratamento (valor- $p=0,047$) e representação emocional (valor- $p=0,037$), difere significativamente entre os 2 grupos (aderentes e não aderentes). De facto, nos três casos, a pontuação média obtida revela-se superior no grupo dos doentes não aderentes à terapêutica ou seja, indivíduos com pontuações mais elevadas revelam-se, potencialmente, com maior risco de serem não aderentes do que aqueles cujas pontuações são mais baixas.

Por fim, verificou-se uma associação significativa entre a dimensão preocupações e a não-adesão. Através do gráfico apresentado na figura 6.18, verificou-se que a pontuação média da subescala é superior para o grupo dos indivíduos não aderentes.

6.3 Análise de Regressão Logística

6.3.1 Análise Preliminar

De forma a fazer uma pré-seleção das variáveis a entrar no modelo de regressão logística, tal como referido no capítulo 3, com base nas variáveis descritas como potenciais



Figura 6.16: Comparação das pontuações (scores) médias entre aderentes e não aderentes para as dimensões identidade, controlo de tratamento e representação emocional, avaliadas pelo (IPQ-R).

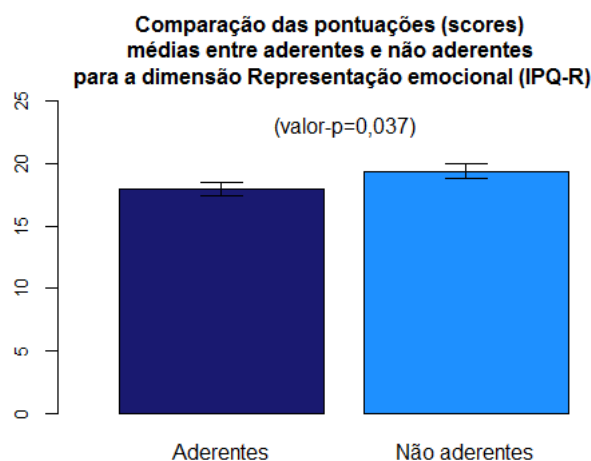


Figura 6.17: Comparação da pontuação (score) média entre aderentes e não aderentes para a dimensão preocupações, avaliada pelo BMQ.

fatores associados à não-adesão e assumindo a não-adesão como uma variável dicotómica, procedeu-se, inicialmente, à análise preliminar através do modelo de regressão logística. Esta análise preliminar consistiu em ajustar a variável resposta em função de cada uma das variáveis preditoras. Com base no valor-p obtido após o ajustamento individual (variável a variável), foram selecionadas para o modelo inicial todas as potenciais variáveis preditoras que mostraram associação significativa com a variável resposta (valor- $p \leq 0.15$) (Paterson *et al.* (2000), Lazo *et al.* (20077) e Fernandes (2011)).

Na tabela 6.26 encontra-se o resultado desta análise preliminar.

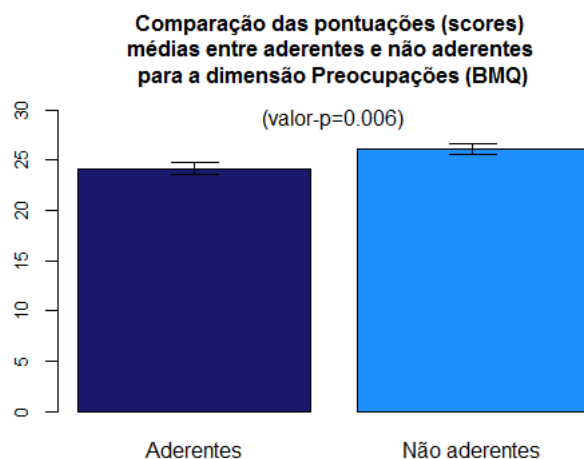


Figura 6.18: Comparação da pontuação (score) média entre aderentes e não aderentes para a dimensão preocupações, avaliada pelo BMQ.

Variável	% aderentes	% não aderentes	valor-p	OR univariado	IC 95
Grupo 1: Variáveis relacionadas com o doente					
Idade (anos)			0,301		
35 ou menos	1,79	4,17		1,00	
Entre 36 e 45	21,43	17,86		0,36	(0,07-1,41)
Mais de 45	30,36	24,40		0,34	(0,07-1,32)
Género			0,812		
Masculino	41,07	36,31		1,00	
Feminino	12,50	10,12		0,92	(0,44-1,89)
País de nascimento			0,449		
Portugal	45,24	37,50		1,00	
País africano	7,74	7,74		1,21	(0,52-2,81)
Outro	0,60	1,19		2,41	(0,23-13,32)
Etnia			0,781		
Branco de origem europeia	48,81	41,67		1,00	
Branco de origem não-europeia	1,79	1,19		0,78	(0,10-4,84)
Negro	1,19	2,38		2,34	(0,44-17,26)
Mulato	1,79	1,19		0,78	(0,10-4,84)
Estado civil			0,666		
Solteiro	22,02	22,62		1,00	
Casado/União de facto	22,62	15,48		0,67	(0,34-1,30)
Divorciado	6,55	5,95		0,89	(0,33-2,34)
Viúvo	2,38	2,38		0,97	(0,22-4,39)
Filhos menores			0,579		
Não	10,71	7,74		1,00	
Sim	42,86	38,69		0,80	(0,36-1,75)
Morar sozinho			0,015		
Não	10,71	17,26		1,00	
Sim	42,86	29,17		2,37	(1,20-4,79)
Número de co-habitantes			0,298		
0-2	44,05	41,07		1,00	

Continua na próxima página

Variável	% aderentes	% não aderentes	valor-p	OR univariado	IC 95%
3	5,36	2,98		0,60	(0,18-1,81)
>4	4,17	2,38		0,61	(0,15-2,12)
Nível de escolaridade			0,988		
Sem habilitações	1,19	0,50		1,00	
1º ciclo	11,31	7,74		1,37	(0,12-12,4)
2º ciclo	6,55	4,17		1,27	(0,10-10,7)
3º ciclo	3,57	10,12		5,67	(0,46-20,6)
Ensino secundário	11,90	8,33		1,40	(0,12-15,4)
Ensino pós-secundário não superior	3,57	4,76		2,67	(0,21-15,4)
Ensino superior	15,48	10,71		1,38	(0,12-12,2)
Situação profissional			0,088		
Desempregado	11,31	7,14		1,00	
Reformado	14,29	8,93		0,99	(0,37-2,63)
Empregado	27,38	28,57		1,65	(0,73-3,86)
Outra	0,60	1,79		4,75	(0,54-24,7)
Horário de trabalho			0,158		
Sem horário	25,00	15,48		1,00	
Fixo	20,24	22,62		1,81	(0,92-3,57)
Por turnos	8,33	8,33		1,62	(0,66-3,96)
Hábitos de exercício físico			0,081		
Sem hábitos	10,71	12,50		1,00	
Ligeiros/moderados	22,02	22,02		0,86	(0,39-1,86)
Intensos	20,83	11,90		0,49	(0,21-1,12)
Hábitos alcoólicos			0,020		
Sem hábitos	26,19	14,29		1,00	
Ligeiros/moderados	10,71	11,31		1,94	(0,86-4,41)
Intensos	16,67	20,83		2,29	(1,14-4,68)
Binge Drinking			<0,001		
Não	48,21	30,95		1,00	
Sim	5,36	15,48		4,50	(2,02-10,87)
Hábitos tabágicos			0,951		
Não	14,88	13,10		1,00	
Sim	38,69	33,33		0,98	(0,50-1,93)
Outros consumos			0,792		
Sim	45,24	39,88		1,00	
Não	8,33	6,55		0,89	(0,37-2,09)
Tratamento sob metadona			0,230		
Não	50,00	45,24		1,00	
Sim	3,57	1,19		0,37	(0,05-1,65)
Satisfação com suporte social	54,0±10,4	52,2±10,9	0,273	0,98	(0,96-1,01)
Ansiedade	2,0±2,7	3,7±4,5	0,006	1,15	(1,05-1,28)
Depressão	2,9±4,0	4,1±4,9	0,087	1,06	(0,99-1,15)
Stress	4,5±4,2	5,4±4,6	0,199	1,05	(0,98-1,13)
Via de transmissão			0,854		
Heterossexuais	26,19	19,05		1,00	
Homens que tiveram sexo com homens /bissexuais	13,10	16,07		1,69	(0,82-3,51)
Utilizadores de drogas injetáveis	7,14	6,55		1,26	(0,49-3,23)
Heterossexuais /Utilizadores de drogas injetáveis	4,17	1,79		0,59	(0,12-2,30)
Outra via de transmissão	2,98	2,98		1,37	(0,36-5,33)

Continua na próxima página

Variável		% aderentes	% não aderentes	valor-p	OR univariado	IC 95%
Grupo 2: Variáveis relacionadas com a situação clínica						
Classificação SIDA				0,118		
	Não	50,0	45,8		1,00	
	Sim	3,57	0,60		0,18	(0,01-1,10)
Presença de co-infecções				0,148		
	Não	38,10	37,50		1,00	
	Sim	15,48	8,93		0,59	(0,28-1,20)
Presença de co-morbilidades				0,723		
	Não	17,86	14,29		1,00	
	Sim	35,71	32,14		0,59	(0,28-1,20)
Situação clínica na 1ª consulta				0,137		
	Assintomático	32,74	32,14		1,00	
	Sintomático	9,52	8,93		0,95	(0,43-2,13)
	Classificação SIDA	11,31	5,36		0,48	(0,19-1,13)
Situação clínica aquando início da TAR				0,244		
	Assintomático	30,36	29,76		1,00	
	Sintomático	11,31	9,52		0,86	(0,39-1,86)
	Classificação SIDA	11,90	7,14		0,61	(0,26-1,37)
Situação clínica na data da entrevista				0,078		
	Assintomático	36,90	36,31		1,00	
	Sintomático	2,98	4,17		1,42	(0,43-5,04)
	Classificação SIDA	13,69	5,95		0,44	(0,19-0,98)
Classificação CDC 1ª consulta				0,248		
	A1	12,50	10,71		1,00	
	A2	14,29	13,69		1,12	(0,48-1,61)
	A3	5,95	7,74		1,52	(0,54-4,36)
	B1	2,98	0,60		0,23	(0,01-1,63)
	B2	2,98	2,98		1,17	(0,28-4,83)
	B3	3,57	5,36		1,75	(0,53-6,13)
	C1	0,00	0,60		2,47	(0,00-14,60)
	C2	1,19	0,60		0,58	(0,03-6,58)
	C3	10,12	4,17		0,48	(0,16-1,38)
Classificação CDC no início da TAR				0,289		
	A1	5,95	4,17		1,00	
	A2	16,07	16,07		1,43	(0,48-4,46)
	A3	7,74	9,52		1,76	(0,53-6,11)
	B1	2,38	0,60		0,36	(0,02-3,11)
	B2	4,76	2,98		0,89	(0,20-3,92)
	B3	4,17	5,95		2,04	(0,53-8,34)
	C1	0,00	0,60		3,03	(0,00-15,76)
	C2	0,60	1,19		2,86	(0,23-12,45)
	C3	11,90	5,36		0,64	(0,18-2,27)
Classificação CDC no momento da entrevista				0,226		
	A1	5,36	2,98		1,00	
	A2	13,69	16,67		2,19	(0,66-7,99)
	A3	8,93	8,33		1,68	(0,46-6,63)
	B1	1,79	0,60		0,60	(0,03-6,29)

Continua na próxima página

Variável		% aderentes	% não aderentes	valor-p	OR univariado	IC 95%
	B2	4,76	2,98		1,12	(0,23-5,52)
	B3	4,17	6,55		2,83	(0,69-12,83)
	C1	0,00	0,60		3,81	(0,00-24,59)
	C1	1,19	1,19		1,80	(0,17-19,26)
	C3	13,69	6,55		0,86	(0,24-3,36)
Efeitos adversos				0,293		
	Não	40,48	32,14		1,00	
	Sim	13,10	14,29		1,37	(0,70-2,73)
Adesão estimada pelo médico				0,404		
	Não aderente	3,57	4,76		1,00	
	Aderente	50,00	41,67		0,63	(0,20-1,88)
Recomendação sobre TAR				0,006		
	Não	29,17	15,48		1,00	
	Sim	24,40	30,95		2,39	(1,28-4,52)
Recomendação de adesão				0,019		
	Não	30,36	17,86		1,00	
	Sim	23,21	28,57		2,09	(1,13-3,91)
Mudança de regime				0,246		
	Não	51,19	42,26		1,00	
	Sim	2,38	4,17		2,12	(0,61-8,36)
Tempo sob TAR				0,022		
	Há 3 meses ou menos	0,60	3,57		1,00	
	Entre 3 e 6 meses	0,60	1,79		0,50	(0,02-15,74)
	Entre 6 e 12 meses	1,19	1,79		0,25	(0,01-3,67)
	Há mais de 12 meses	51,19	39,29		0,13	(0,01-0,77)
Grupo 3: Variáveis relacionadas com os serviços de saúde e medicação/doença						
Identidade		2,83±3,14	4,26±3,84	0,011	1,12	(1,03-1,23)
Duração aguda/crônica		23,91±3,46	23,60±4,24	0,602	0,98	(0,90-1,06)
Duração cíclica		9,87±3,09	10,53±3,88	0,222	1,06	(0,97-1,16)
Consequências		20,18±4,92	21,18±5,02	0,194	1,04	(0,98-1,11)
Controlo pessoal		22,84±3,49	22,81±4,02	0,949	1,00	(0,92-1,08)
Controlo de tratamento		18,21±2,85	19,05±3,16	0,074	1,10	(0,99-1,22)
Coerência da doença		11,76±3,34	12,01±3,81	0,639	1,02	(0,94-1,11)
Representação emocional		17,97±4,79	19,37±5,21	0,072	1,06	(1,00-1,13)
Preocupações		24,18±5,14	26,13±4,65	0,013	1,09	(1,02-1,16)
Necessidades		32,66±3,68	32,42±3,42	0,672	0,98	(0,90-1,07)

Tabela 6.26: Análise univariada dos fatores da não-adesão

6.3.2 Modelo de Regressão Logística Múltipla

Após a aplicação do modelo de regressão logística múltipla com seleção por *stepwise* verificou-se que as variáveis morar sozinho, *binge drinking*, ansiedade, estado clínico no momento da consulta, controlo de tratamento e preocupações com os fármacos, mostraram-se independentemente associadas à não-adesão.

De acordo com a tabela 6.27, comportaram-se como fatores de risco para a não-adesão,

o facto do indivíduo morar sozinho, ser *binge drinker*, apresentar sintomas de ansiedade, manifestar controlo sobre o tratamento e apresentar preocupações sobre os efeitos secundários dos fármacos. Por outro lado, o facto do indivíduo possuir classificação SIDA no momento da aplicação do questionário, mostrou-se fator de proteção de não-adesão.

Tabela 6.27: Análise de regressão logística multivariada, com *stepwise*, dos fatores associados à não-adesão

Variável	β	OR ¹	IC ² 95%	valor-p
(Intercept)	-7,51	0,01	(0,00 – 0,11)	< 0,01
Morar sozinho				
Não				
sim	1,17	3,23	(0,14 – 7,73)	0,006
Binge Drinking				
Não				
sim	1,43	4,18	(1,71 – 11,02)	0,002
Ansiedade (EADS)	0,15	1,16	(1,03 – 1,34)	0,025
Situação clínica no momento da entrevista				
Assintomático				
Sintomático	-0,19	0,83	(0,20 – 3,52)	0,793
Classificação SIDA	-1,18	0,31	(0,10 – 0,82)	0,023
Identidade (IPQ-R)	0,10	1,10	(0,97 – 1,25)	0,139
Controlo de tratamento (IPQ-R)	0,22	1,25	(1,10 – 1,44)	0,001
Preocupações (BMQ)	0,08	1,09	(1,00 – 1,18)	0,036

¹OR - Estimativa *odds ratio* ajustado. ²IC - Intervalo de confiança.

6.3.2.1 Diagnóstico do Modelo

De acordo com o que foi referido na secção 4.2.5, um dos pressupostos inerentes ao modelo de regressão logística é de que os resíduos não apresentem um padrão e que, pelo menos, 95% estejam contidos no intervalo $[-2,2]$, ou por outras palavras, que não existam muitos resíduos que sejam maiores que 2 em valor absoluto. A partir deste limite, os resíduos são considerados elevados. Com base na figura 6.19, na qual se apresentam os gráficos dos resíduos correntes (ou da Deviance), dos resíduos do qui-quadrado de Pearson e dos resíduos padronizados, verifica-se que este pressuposto é verificado. Observa-se-se que os resíduos não apresentam padrão algum, pelo que não existe evidências para assumir a heterocedasticidade. Por fim, através dos gráficos é, ainda, visível uma grande concentração dos resíduos em torno do valor 0, com uma amplitude constante, o que se traduz na não existência de anomalias no ajustamento do modelo, como por exemplo, função de ligação errada.

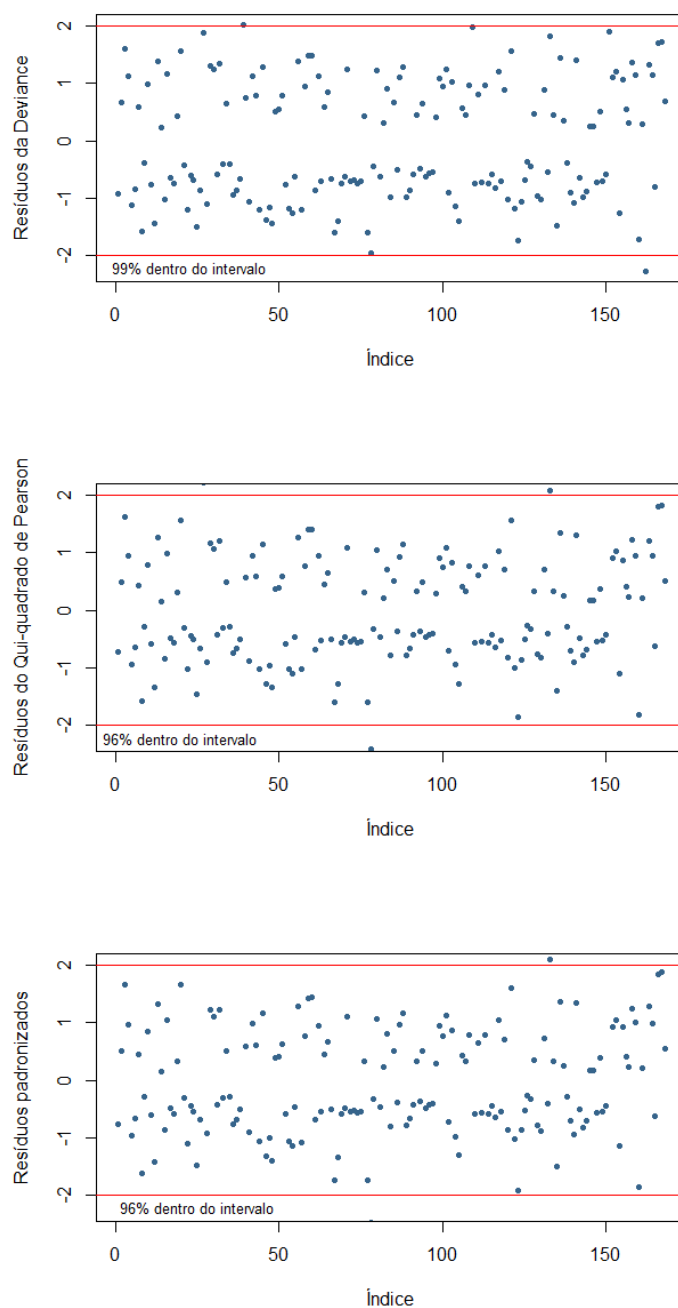


Figura 6.19: Gráficos dos resíduos da deviance, dos resíduos do qui-quadrado de Pearson e dos resíduos padronizados

Na figura 6.20 encontram-se os gráficos dos resíduos (Deviance, qui-quadrado de Pearson e padronizados) *vs* valores ajustados. Através desta figura é possível verificar que não existem evidências de um mau ajustamento. A correlação entre os valores ajustados e os resíduos do modelo é 0.13, valor este que se encontra próximo de 0.

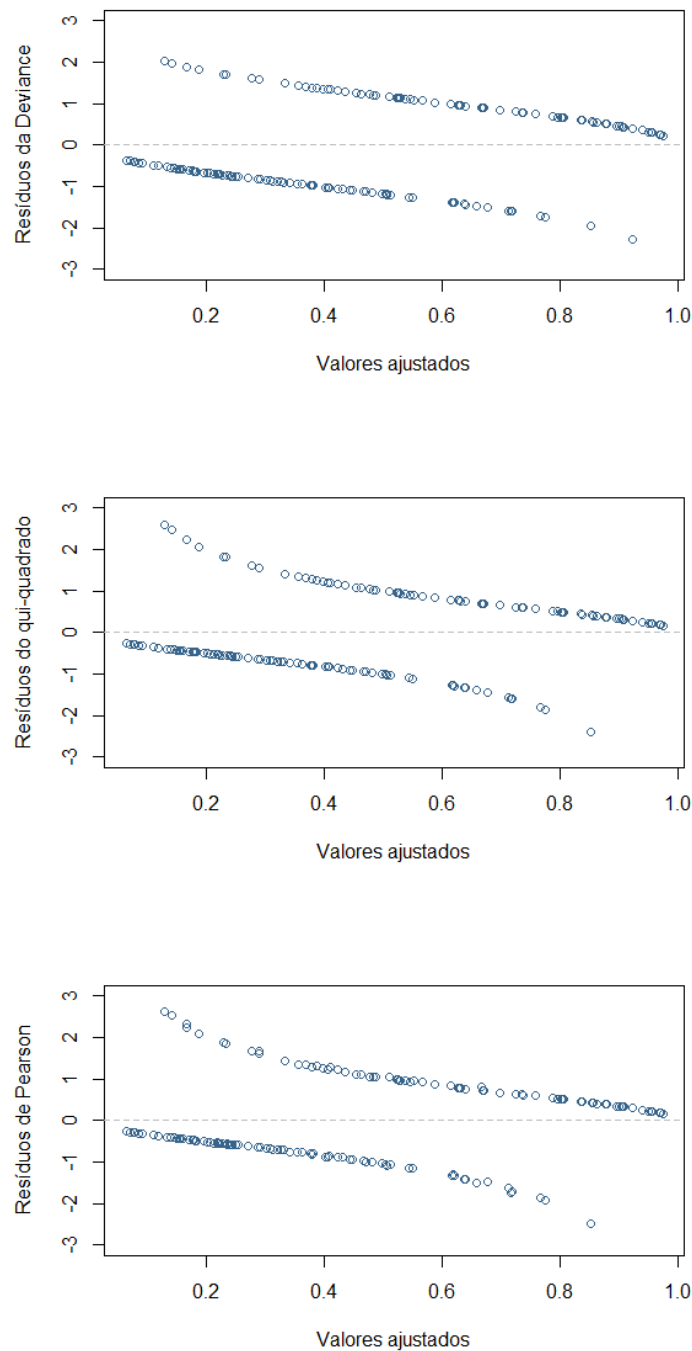


Figura 6.20: Gráficos dos resíduos da deviance *vs* valores ajustados, dos resíduos do qui-quadrado de Pearson *vs* valores ajustados e dos resíduos padronizados *vs* valores ajustados

Ao contrário do que acontece com o modelo de regressão linear (simples ou múltipla), o modelo de regressão logística, e mais, geralmente, o modelo linear generalizado, não exige a normalidade dos resíduos. No entanto, a representação do QQ-plot Normal (figura 6.21) é útil na verificação da estabilidade da variância dos resíduos. Para além disso, o gráfico

apresentado na figura 6.21, evidencia a existência de dois grupos distintos de observações, motivado pelo facto da variável resposta ser binária, o que conduz a que os resíduos sejam constituídos por dois grandes grupos - o negativo e o positivo, não existindo valores nulos.

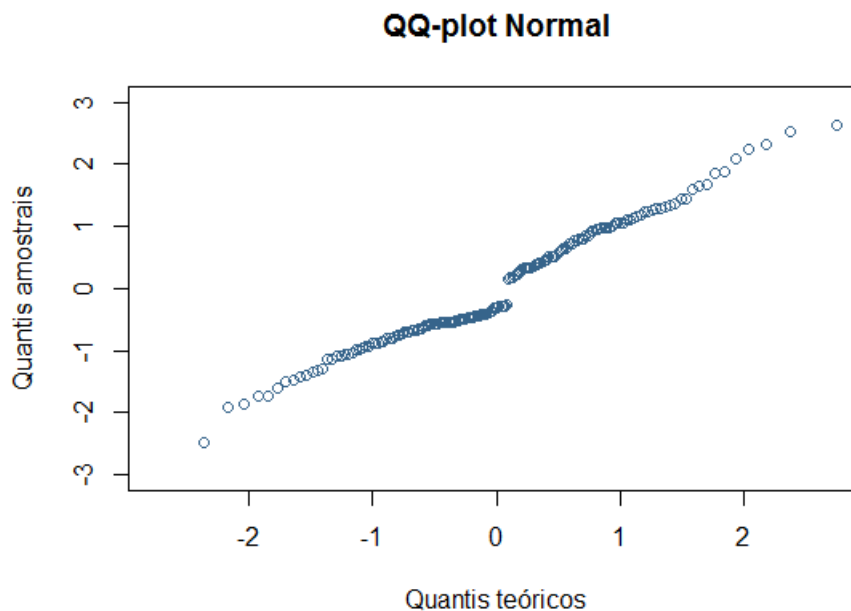


Figura 6.21: QQ-plot Normal dos resíduos do modelo

Segundo o teste de Hosmer-Lemeshow, o modelo final (tabela 6.27) foi considerado adequado ($X^2_{HL_{obs}} = 1,611$, valor- $p=0,991$).

6.3.2.2 Capacidade Preditiva do Modelo

Uma vez analisados os resíduos do modelo, é necessário ainda, avaliar a capacidade preditiva do modelo, isto é, analisar o ajuste do modelo aos dados em estudo. Esta análise é efetuada com recurso à curva ROC, pelo que na figura 6.22 apresenta-se a curva ROC subjacente ao modelo final. De acordo com o método referido na secção 4.2.7 (Método de *Yonden's Index*), encontram-se traçadas, a vermelho, a reta que traduz a aleatoriedade, a cinzento, uma paralela à reta que traduz a aleatoriedade e, por fim, a verde, uma outra paralela à reta a vermelho e a qual é tangente à curva ROC. Através desta última reta, determina-se o par (sensibilidade, 1-especificidade) ótimo. Desta forma, a área sob a curva ROC foi de 0,797, com uma sensibilidade de 81,2% e uma especificidade de 65,6%.

Na figura 6.23 é possível verificar a evolução dos acertos, sensibilidade, especificidade e desempenho do modelo para os diferentes pontos de corte possíveis. A verde encontra-

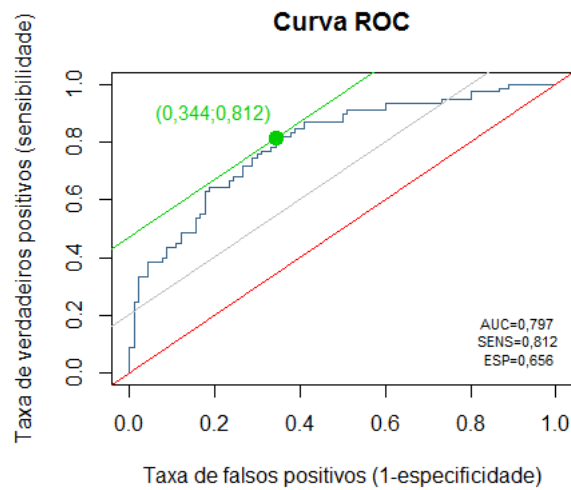


Figura 6.22: Curva ROC do modelo

se assinalado o ponto de corte ótimo (0,384), ao qual corresponde uma sensibilidade de 81,2% e uma especificidade de 65,6%.

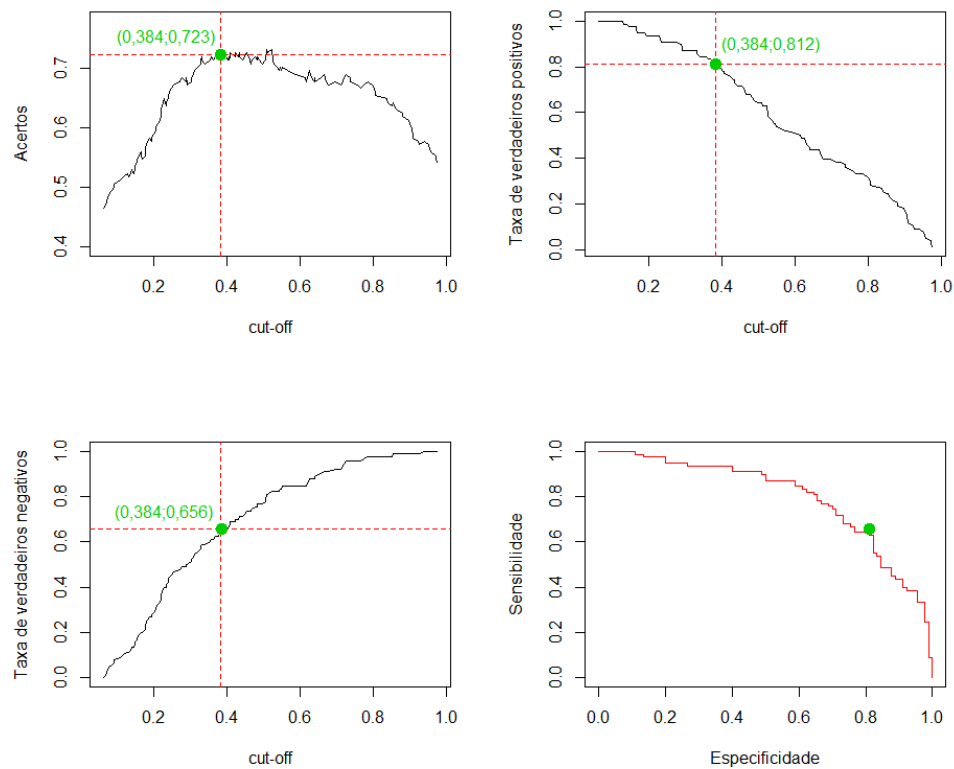


Figura 6.23: Acertos, sensibilidade, especificidade e desempenho do modelo

6.4 Análise de Correspondências Múltiplas

Com o objetivo de traçar os perfis dos doentes aderentes e não aderentes utilizou-se, ainda, a técnica multivariada denominada de Análise de Correspondências Múltiplas.

Para o algoritmo de ACM foram, inicialmente, consideradas as variáveis que foram submetidas ao modelo inicial de regressão logística múltipla.

Tal como foi mencionado no capítulo 5, esta abordagem é aplicada sobre variáveis categóricas, pelo que foi necessário, em primeiro lugar, categorizar as variáveis número de co-habitantes, depressão, ansiedade, identidade, controlo de tratamento, representação emocional e preocupações.

É importante referir que a categorização das variáveis acima indicadas origina perda de informação, pelo que os resultados obtidos com as duas metodologias não têm que ser, obrigatoriamente iguais. Todavia, espera-se que os fatores associados à não-adesão com as duas abordagens sejam semelhantes.

6.4.1 Categorização das Variáveis de Interesse

Para as subescalas de ansiedade e depressão obtiveram-se as seguintes características amostrais:

```
> desc_EADS
```

	Ansiedade	Depressão
Mínimo	0.0	0
1º Quartil	0.0	0
Mediana	2.0	2
3º Quartil	3.5	5
Máximo	21.0	21

De acordo com os quartis apresentados acima, consideraram-se, para a subescala de ansiedade, as seguintes categorias: $[0,3[$, $[3,5[$ e $[5,21[$. No que respeita à subescala de depressão, para a aplicação da ACM foram consideradas as seguintes categorias: $[0,3[$, $[3,6[$ e $[6,21[$.

Na figura 6.24, encontram-se os gráficos de barras das categorias consideradas para cada uma das duas subescalas.

No que respeita às três dimensões do *IPQ-R*, apresentam-se abaixo as características amostrais utilizadas para a categorização das mesmas:

```
> desc_IPQ
```

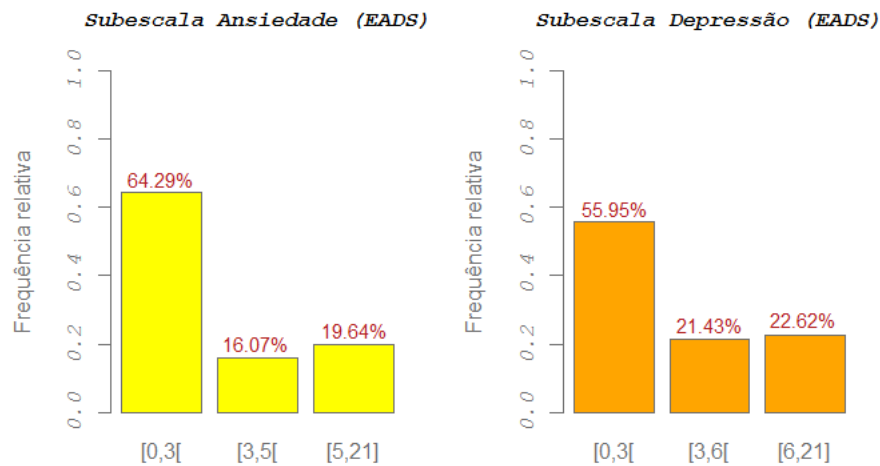


Figura 6.24: Gráfico de barras das categorias das subescalas (variáveis) ansiedade e depressão.

	I	CT	RE
Mínimo	0	11	6
1º Quartil	0	16	14
Mediana	2	18	18
3º Quartil	6	21	23
Máximo	14	25	0

Para a subescala de identidade foram consideradas as categorias [0,3[, [3,7[e [7,14]. Por sua vez, para a subescala controlo de tratamento, consideraram-se as categorias [11,17[, [17,19[, [19,22[e, finalmente, para a subescala de representação emocional, foram utilizadas as seguintes categorias: [6,15[, [15,19[, [19,24[e [24,30].

Na figura 6.25 apresentam-se os gráficos de barras para as categorias de cada uma das subescalas do instrumento *IPQ-R*.

Por fim, a dimensão de preocupações, avaliada pelo instrumento BMQ obtiveram-se as seguintes características amostrais:

```
> desc_BMQ
```

	Preocupações
Mínimo	10
1º Quartil	22
Mediana	25
3º Quartil	28
Máximo	38

Desta forma, para a subescala de preocupações foram consideradas as categorias [10,23[, [23,26[, [26,29[, [29,39].

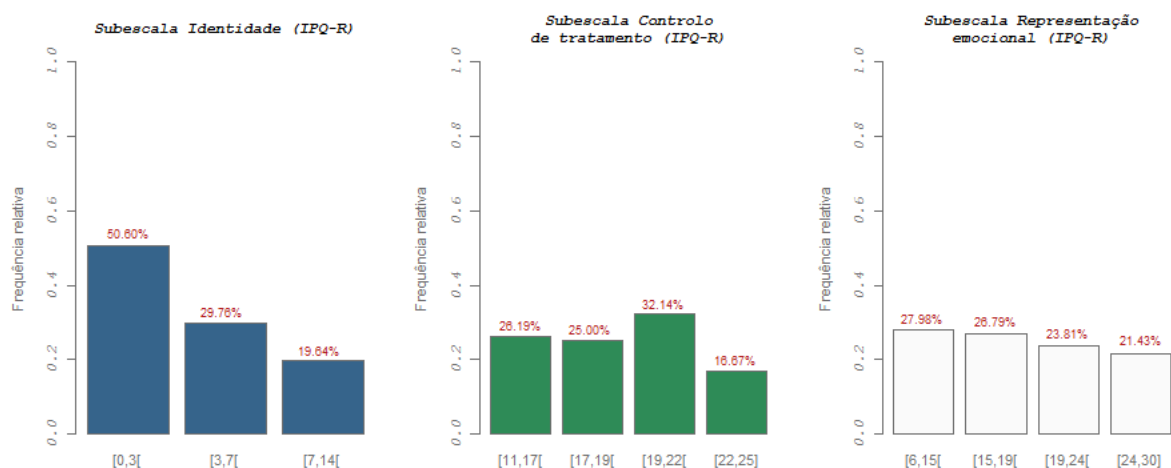


Figura 6.25: Gráfico de barras das categorias das subescalas (variáveis) identidade, controlo de tratamento e representação emocional, respetivamente.

Na figura 6.26 encontra-se os gráficos de barras para a dimensão de preocupações de acordo com as categorias consideradas.

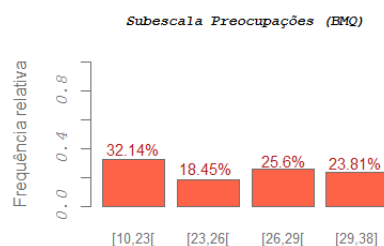


Figura 6.26: Gráfico de barras das categorias da subescala (variável) preocupações.

6.4.2 Aplicação da ACM

Como foi referido anteriormente, inicialmente, de forma a caracterizar os diferentes perfis dos doentes aderentes e não aderentes, foi aplicada a metodologia de ACM sobre as variáveis que se mostraram significativamente associadas à variável resposta aquando a análise preliminar efetuada na secção anterior.

De seguida, consideraram-se apenas as variáveis que permaneceram no modelo final de regressão logística múltipla de forma a verificar o comportamento das mesmas.

Com vista a uma representação gráfica mais perceptível foi necessário, em primeiro lugar, renomear as variáveis, pelo que, de seguida, se apresenta um resumo das mesmas:

AACTG:	Situação clínica na 1 ^a consulta:
AACTG0→ aderente	SCI1→ Assintomático
AACTG1→ não aderente	SCI2→ Sintomático
Morar sozinho:	SCI2→ Classificação SIDA
MS1→ Sim	Situação clínica aquando entrevista:
MS2→ Não	SCE1→ Assintomático
Situação profissional:	SCE2→ Sintomático
SP1→ Desempregado	SCE2→ Classificação SIDA
SP2→ Reformado	Recomendação de adesão:
SP3→ Empregado	RA1→ Sim
SP4→ Outra	RA2→ Não
Prática de exercício físico:	Tempo sob TAR:
EX1→ Sem hábitos	TTAR2→ Há 3 meses ou menos
EX2→ Hábitos ligeiros/moderados	TTAR3→ Há 6 meses ou menos
EX3→ Hábitos intensos	TTAR4→ Há 12 meses ou menos
Hábitos alcoólicos:	TTAR5→ Há mais de 12 meses
ALC1→ Sem hábitos	Identidade (<i>IPQ-R</i>):
ALC2→ Hábitos ligeiros/moderados	IPQI1→ [0,3[
ALC3→ Hábitos intensos	IPQI2→ [3,7[
<i>Binge drinking</i> :	IPQI3→ [7,14[
BD1→ Sim	Controlo de tratamento (<i>IPQ-R</i>):
BD2→ Não	IPQCT1→ [11,17[
Ansiedade (EADS):	IPQCT2→ [17,19[
EADSA1→ [0,3[IPQCT3→ [19,22[
EADSA2→ [3,6[IPQCT4→ [22,25[
EADSA3→ [6,21[Representação emocional (<i>IPQ-R</i>):
Depressão (EADS):	IPQRE1→ [6,15[
EADSD1→ [0,3[IPQRE2→ [15,19[
EADSD2→ [3,5[IPQRE3→ [19,24[
EADSD3→ [5,21[IPQRE4→ [24,30[
Classificação SIDA:	Preocupações (BMQ):
SIDA1→ Sim	BMQP1→ [10,23[
SIDA2→ Não	BMQP2→ [23,26[
Presença de co-infecções:	BMQP3→ [26,29[
COI1→ Sim	BMQP4→ [29,38[
COI2→ Não	

Após a aplicação da ACM sobre os 177 participantes, obtém-se o mapa apresentado na figura 6.27.

A ACM foi efetuada através da matriz disjuntiva completa, sendo, assim, possível apresentar os indivíduos através de pequenos pontos a preto. Por sua vez, as categorias

É possível verificar, por exemplo, através de uma observação rápida do gráfico apresentado na figura 6.27, que parece existir uma relação forte entre o facto do participante se encontrar sobre tratamento há três meses ou menos (TTAR2) e o facto de ser *binge drinker* (BD1).

Tendo em conta que na parte central do gráfico existe um aglomerado de categorias, tornando o gráfico de difícil interpretação, nas figuras 6.29 e 6.28, apresenta-se o mesmo gráfico com *zoom* em redor das categorias dos indivíduos aderentes (AACTG0) e não aderentes (AACTG1), respetivamente.

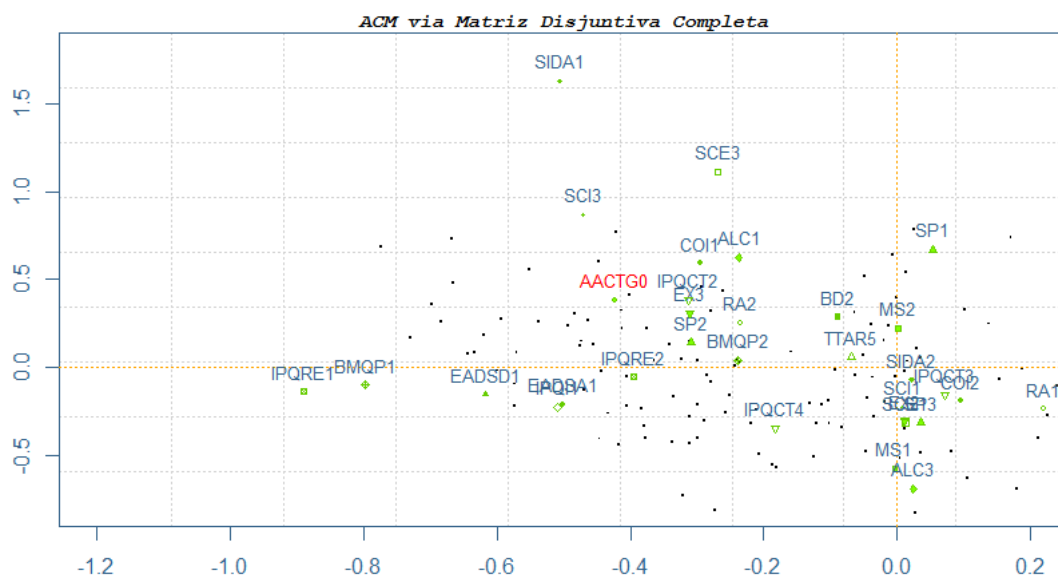


Figura 6.28: Representação gráfica das duas primeiras dimensões da análise de correspondências múltiplas com *zoom* em redor da categoria referente aos indivíduos aderentes (AACTG0)

Com base na figura 6.28, é possível verificar, por exemplo, que existe uma forte relação entre indivíduos que não possuíam sintomatologia SIDA (SIDA2), que não manifestavam co-infecções (COI2), que praticavam exercício esporadicamente (EX2), que não apresentavam sintomas da doença na 1ª consulta (SCI1) e que manifestavam pouco controlo sobre o tratamento.

Em relação ao gráfico 6.29, observa-se, por exemplo, que existe uma relação forte entre indivíduos com estados depressivos moderados/elevados (EADSD3) e participantes que apresentam pontuações elevadas para a dimensão representação emocional (IPQRE4), isto é, que se mostram perturbados, amedontrados e/ou ansiosos quando pensam na doença.

No entanto, de forma a ser possível aferir, de forma mais clara, as relações mais

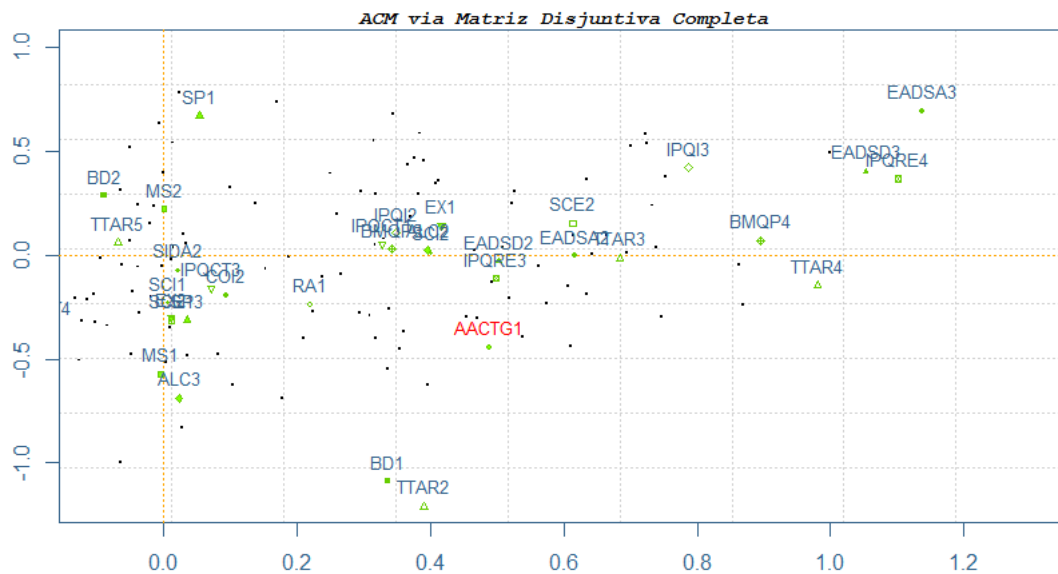


Figura 6.29: Representação gráfica das duas primeiras dimensões da análise de correspondências múltiplas com *zoom* em redor da categoria referente aos indivíduos não aderentes (AACTG1)

significativas, na figura 6.30 é apresentado o dendrograma, obtido pelo método da ligação completa, com base na matriz de dissimilaridades calculada a partir das coordenadas das categorias no gráfico apresentado na figura 6.27. O dendrograma permite, através de uma representação gráfica simples, extrair as relações mais fortes. Para além disso, permite verificar a ordem de grandeza da relação entre categorias.

Desta forma, na figura 6.31 reproduz-se o mapa apresentado na figura 6.27 com os grupos sugeridos pelo dendrograma.

Com base nos gráficos 6.30 e 6.31, observam-se vários grupos que se podem caracterizar da seguinte forma:

- Indivíduos que mostram pouco preocupados, zangados com a doença (IPQRE1 e IPQRE2), que não demonstram preocupação relativamente aos fármacos (BMQP1), que, praticamente, não associam os sintomas à doença (IPQI1) e doentes que não revelam sintomas ou revelam fracos sintomas de ansiedade e depressão (EADSA1 e EADSD1, respetivamente) (a vermelho).

- Indivíduos aos quais não foi recomendada a adesão à terapêutica (RA2), que se encontram reformados (SP2), que mostram algumas preocupações acerca dos fármacos (BMQP2), que se encontram sob TAR há mais de doze meses (TTAR5), que não moram sozinhos (MS2) e que não são *binge drinkers* (BD2)(a verde).

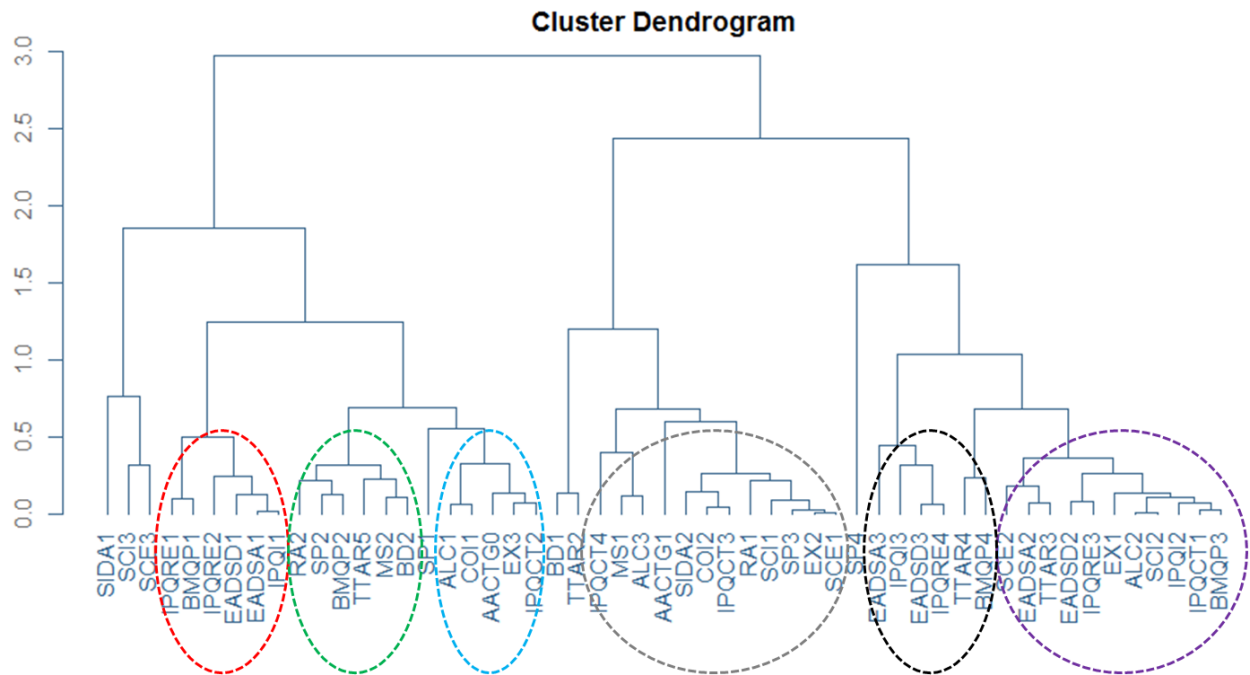


Figura 6.30: Dendrograma das coordenadas das duas primeiras dimensões da análise de correspondências múltiplas cujo gráfico se apresenta na figura 6.27

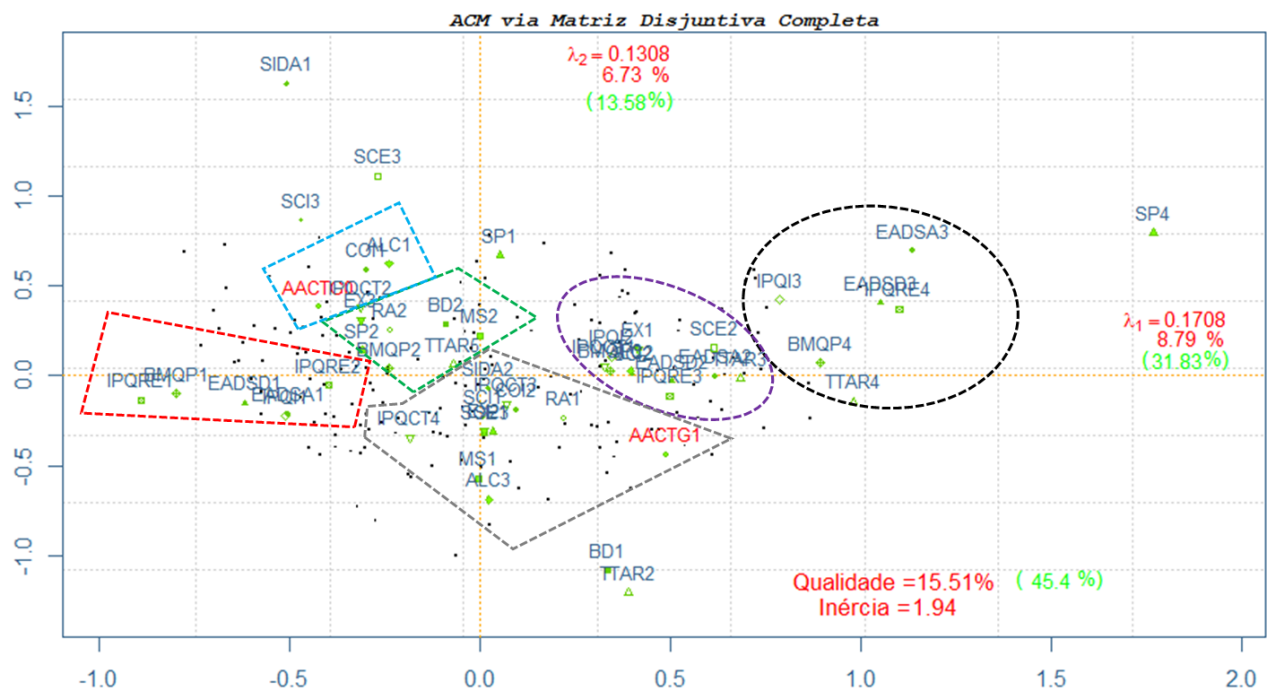


Figura 6.31: Replicação do gráfico resultante da ACM.

- Indivíduos que se encontram sob TAR entre seis a doze meses (TTAR4), que demonstram elevada preocupação e revolta com a doença (IPQRE4), que revelam grande

preocupação em relação aos fármacos (BMQP4), com sintomas elevados de depressão (EADSD3) e ansiedade (EADSA3) e que associam grande parte dos sintomas à doença (IPQI3) (a preto).

Note-se que vários estudos concluíram que a ansiedade e a depressão tendem a estar fortemente relacionadas, pelo que se desenvolveu o modelo tripartido (Watson *et al.*, 1995), isto é, o modelo composto por três dimensões: ansiedade, depressão e stress. Esta relação é igualmente verificada nos gráficos acima, uma vez que, cada uma das três categorias da dimensão ansiedade, encontra-se relacionada com a respetiva categoria da subescala depressão.

Para além dos grupos acima indicados, é possível identificar os grupos nos quais se encontram incluídas as categorias referentes aos doentes não aderentes (ACCTG1) e aos aderentes (AACTG0) e, como tal, estabelecer os perfis dos indivíduos aderentes e não aderentes:

- **Doentes aderentes:**

Indivíduos que não possuem hábitos alcoólicos (ALC1), que manifestam co-infecções (COI1), que praticam exercício físico regularmente (EX3) e que consideram o tratamento pouco importante (IPQCT2) (a azul).

- **Doentes não aderentes:**

Indivíduos que moram sozinhos (MS1), que possuem hábitos alcoólicos intensos (ALC3), aos quais foi efetuada recomendação de adesão pelo médico assistente (RA1), que consideram que o tratamento tem uma importância moderada/elevada para o controlo da doença (IPQCT3 e IPQCT4), que não possuem sintomatologia SIDA (SIDA2), que não apresentam co-infecções (COI2), que se encontram empregados (SP3), que não apresentavam sintomas da doença na 1^a consulta (SCI1) e na consulta efetuada no momento em que foram convidados a participar no estudo (SCE1) e que praticam exercício esporadicamente (EX2) (a cinzento).

Uma outra abordagem via ACM consistiu em considerar apenas as variáveis incluídas no modelo final, isto é, as variáveis que se encontram na tabela 6.27.

Desta forma, após aplicação da ACM via matriz disjuntiva completa, obteve-se o gráfico apresentado na figura 6.32. Utilizando as coordenadas das modalidades, as quais retratam valores de dissemelhança, obteve-se o dendrograma de apoio à ACM, ilustrado na figura 6.33.

De acordo com o mapa da ACM, observa-se que o primeiro eixo explica, aproximadamente, 38,81% da variação total dos dados, enquanto o segundo eixo explica, sensi-

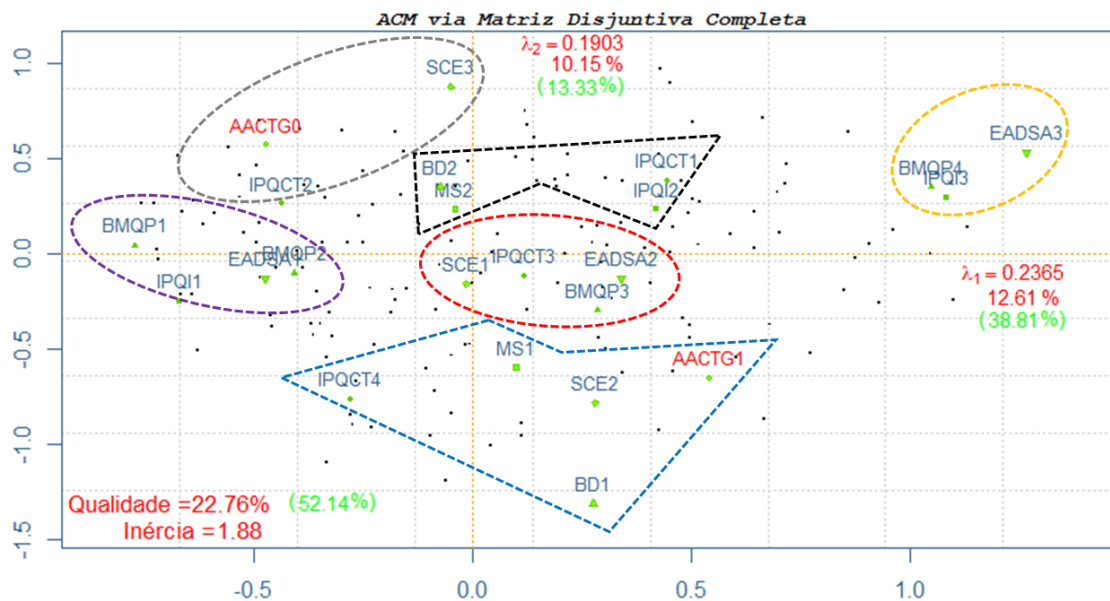


Figura 6.32: Representação gráfica das duas primeiras dimensões da análise de correspondências múltiplas considerando as variáveis incluídas no modelo final de regressão logística múltipla

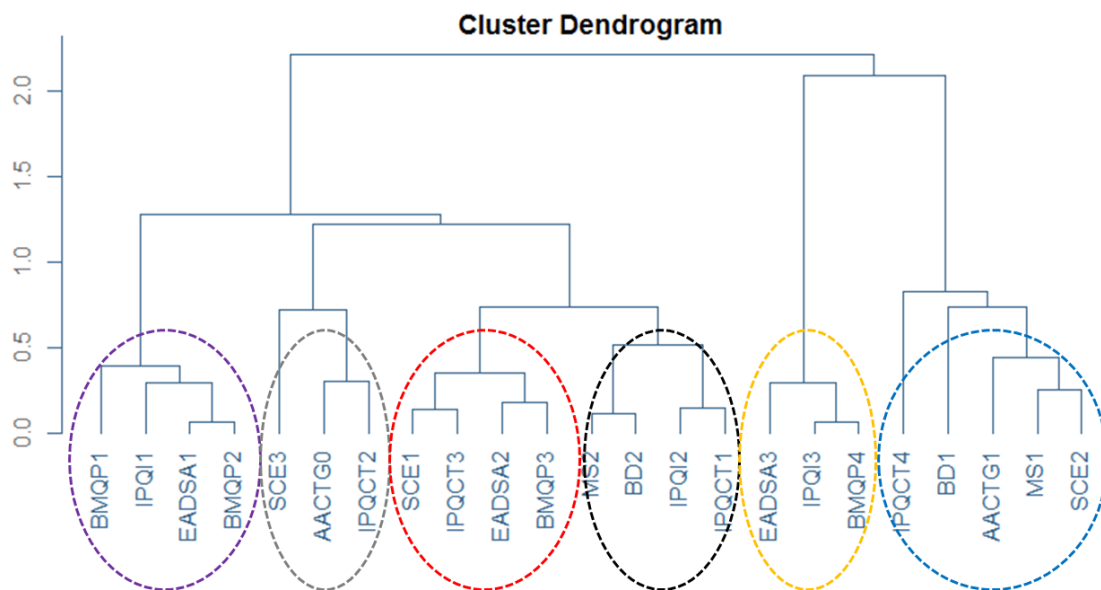


Figura 6.33: Dendrograma de apoio ao resultado da ACM ilustrado no gráfico 6.32

velmente, 13,33%. Desta forma, a solução obtida explica, aproximadamente, 52,14% da variabilidade total dos dados.

Através de uma observação rápida do gráfico resultante da ACM e o qual se encontra apresentado na figura 6.32, verifica-se que sintomas inexistentes ou fracos de ansiedade (EADSA1) se encontram fortemente relacionados com preocupações fracas a moderadas

sobre os fármacos anti-retrovíricos (BMQP2). Verifica-se uma situação semelhante no que diz respeito às categorias associadas a preocupações elevadas com os medicamentos (BMQP4) e o facto de associar grande parte dos sintomas à doença (IPQI3).

Para além disso, observa-se, também, que as categorias relativas a pontuações elevadas de ansiedade (EADSA3), identidade (IPQI3) e preocupações com os medicamentos (BMQP4), apesar de não parecerem relacionadas com as demais variáveis, evidenciam uma forte relação entre si.

Com base nos gráficos apresentados nas figuras 6.32 e 6.33, observam-se vários grupos que se podem caracterizar da seguinte forma:

- Indivíduos que manifestam fraca a moderada preocupação com os medicamentos anti-retrovíricos (BMQP1 e BMQP2), que, praticamente, não associam os sintomas à doença (IPQI1) e que não revelam indícios de ansiedade (EADSA1) (a roxo).

- Indivíduos que não apresentavam sintomas aquando a consulta efetuada no momento da entrevista (SCE1), que manifestam ligeiros a moderados sintomas de ansiedade (EADSA2), que revelam preocupação moderada/elevada sobre os fármacos (BMQP3) e que consideram o tratamento com uma importância moderada/elevada para o controlo da doença (IPQCT3) (a vermelho).

- Indivíduos que associam grande parte dos sintomas à doença (IPQI3), que revelam uma preocupação elevada com os medicamentos anti-retrovíricos e com os seus efeitos a médio e/ou longo prazo (BMQP4), que consideram o tratamento altamente importante para o controlo da doença (IPQCT4) e que apresentam elevados indícios de ansiedade (EADS3) (a laranja).

- Indivíduos que não moram sozinhos (MS2), que não são *binge drinkers* (BD2), que atribuem pouco importância ao tratamento (IPQCT1) e que associam poucos sintomas à doença (IPQI2) (a preto).

À semelhança da abordagem aplicada anteriormente, na qual se consideraram as variáveis submetidas ao modelo inicial de regressão logística múltipla, para além dos grupos acima indicados, é possível estabelecer os grupos nos quais se encontram incluídas as categorias referentes aos doentes não aderentes (ACCTG1) e aos aderentes (AACTG0) e, como tal, identificar os perfis dos indivíduos aderentes e não aderentes:

- **Doentes aderentes:**

Indivíduos que possuem fraco a moderado controlo do tratamento (BMQP2) e que apresentavam sintomatologia SIDA no momento da entrevista (SCE3) (a cinzento).

- **Doentes não aderentes:**

Indivíduos que são *binge drinkers* (BD1), que moram sozinhos (MS1), que consideram o tratamento bastante importante para o controlo da doença (IQPCT4) e que são sintomáticos (B) (SCE2) (a azul).

Capítulo 7

Discussão e Problemas em Aberto

A não-adesão é um problema frequente e deve ser esperado em todos os doentes que recebem tratamento anti-retrovírico. No presente estudo, verificou-se que 48,6% dos participantes se revelaram não aderentes à terapêutica anti-retrovírica. Estudos que investigaram a adesão ao tratamento anti-retrovírico em adultos encontraram prevalência de não-adesão próximas da encontrada no presente trabalho (Nieuwkerk *et al.* (2000), Marin-Sanchez *et al.* (2000)).

O primeiro passo para prevenir esta situação consiste na identificação precisa, tanto quanto possível, dos doentes cujo risco de não-adesão é suficiente para enfraquecer a resposta imunológica resultando no menor controlo da infeção e piores resultados clínicos.

A adesão à terapêutica é "um problema de etiologia multifatorial" (European Centre for Disease Prevention and Control/WHO Regional Office for Europe, 2010) e depende de um determinado ambiente clínico e cultural. Os resultados obtidos com o presente estudo suportam estudos anteriores, identificando vários fatores associados à não-adesão à TARc, de entre as variáveis relacionadas com o doente, a sua situação clínica e a relação do doente com os serviços de saúde (European Centre for Disease Prevention and Control/WHO Regional Office for Europe (2010), Jordan *et al.* (2002)). A correta identificação dos fatores permitirá o estabelecimento de estratégias futuras mais efetivas, que tendem a basear-se na adaptação do regime terapêutico ao estilo de vida do doente e em avaliar a adesão como parte de um programa de seguimento.

Fatores relacionados com o doente

A associação entre fatores socio-demográficos e a não-adesão à TARc não é consensual (Yach (2003), Chesney (2000)). No presente estudo verificou-se que, à semelhança de estudos anteriores, o estado civil, a idade e o género não se revelaram significativamente associados à não-adesão (Mitiku *et al.* (2013), Peltzer *et al.* (2010), Talam *et al.* (2008),

Campbell *et al.* (2010)). De referir, no entanto, que, embora a variável idade não se tenha mostrado estatisticamente associada à não-adesão, foi encontrada uma menor adesão entre os doentes com idade compreendida entre os 18 e 35 anos.

Em grande parte dos estudos verifica-se que o nível de escolaridade encontra-se associado à não-adesão, onde indivíduos com um nível mais baixo de escolaridade são mais propícios a ser não aderentes (Blatt *et al.* (2009), Carvalho *et al.* (2003), Lignani Júnior *et al.* (2001), Melchior *et al.* (2007), Nemes *et al.* (2000)). Esta associação pode estar relacionada com a dificuldade de entendimento, por parte dos doentes com nível escolar mais baixo, das instruções de toma. Contudo, no presente estudo não foi verificada tal associação. Tal facto, poder-se-á dever a vários fatores externos ao doente, tal como, um bom acompanhamento por parte dos profissionais de saúde.

De entre as variáveis relacionadas com o doente, o facto do paciente morar ou não sozinho mostrou-se significativamente associado à não-adesão, mantendo-se após o ajustamento das restantes variáveis (valor- $p=0,006$). Mais concretamente, verificou-se que o facto do paciente morar sozinho é fator de risco da não-adesão. Esta conclusão não é estranha, uma vez que o apoio da família e amigos é fundamental para a promoção de uma boa adesão. Desta forma, as pessoas que partilham a habitação com o doente poderão lembrar o mesmo do horário da toma da medicação, contribuindo, assim, para o cumprimento do esquema terapêutico. São vários os estudos que suportam este resultado (Carvalho *et al.* (2003)). No entanto, é de realçar que, quanto a este fator, não há consenso na literatura. Em estudos anteriores como o de Riera *et al.* (1999) registou-se uma menor adesão entre os que vivem com os progenitores. Vale (1999) encontrou uma menor adesão em indivíduos que vivem em instituições ou moram sozinhos.

Das características de ansiedade, depressão e stress, foi observado que o nível de ansiedade se revelou significativamente associado à não-adesão, mesmo após ser submetida à análise de regressão múltipla (valor- $p=0,025$). Indivíduos com níveis mais elevados de ansiedade mostraram-se propícios à não-adesão. Este resultado é suportado por vários estudos efetuados na área. Ammassari *et al.* (2002) avaliaram a associação entre a ansiedade como medida do Índice Mental - 5, através do auto-relato de adesão de uma coorte diversificada de 358 italianos que viviam com HIV. Após a análise multivariada, foi descoberto que os sintomas de ansiedade foram independentemente associados à não adesão aos fármacos anti-retrovirais nos últimos 3 dias anteriores ao questionário. Da mesma forma, Van Servellen *et al.* (2002) concluíram que o score de ansiedade foi significativamente associado à não-adesão por auto-relato, bem como à adesão avaliada pelo médico com base nos registos clínicos. Campos *et al.* (2010) investigaram os sintomas de ansiedade e depressão como fatores de risco para a não-adesão aos medicamentos anti-retrovíricos

numa amostra composta por 293 doentes a iniciar a terapêutica em dois centros de referência no Brasil. A análise bivariada revelou um aumento do risco de não-adesão para os doentes com sintomas graves de ansiedade. Após o ajustamento dos restantes fatores, observou-se que os sintomas graves de ansiedade foram independentemente associados à não-adesão.

No presente trabalho, a depressão e o stress não se revelaram independentemente associados à não-adesão, isto é, não foram incluídas no modelo final. No entanto, é importante referir que, apesar da depressão não se revelar independentemente associada à não-adesão, entrou no modelo inicial. Tal como no presente estudo, no trabalho realizado por Campos *et al.* (2010), não houve indicação de associação estatística entre a depressão grave e a não-adesão.

Ainda em relação aos fatores associados ao doente, à semelhança de vários estudos realizados, verificou-se que o consumo de álcool, em particular, o consumo excessivo de álcool num curto espaço de tempo, foi independentemente associado à não-adesão (valor- $p=0,002$). Indivíduos considerados *binge drinkers* possuem maior propensão à não-adesão quando comparados com indivíduos sem hábitos alcoólicos. O consumo de álcool é um fator de risco para a baixa adesão à TAR (Cook *et al.* (2001), Lucas *et al.* (2002), Samet *et al.* (2004) e Tucker *et al.* (2003)). De acordo com Braithwaite and Kendall (2010) o consumo de álcool pode conduzir à redução da adesão à TAR, levando a uma diminuição da eficácia da terapêutica e, finalmente, a um aumento da mortalidade relacionada com o VIH. Para além disso, vários estudos registaram o impacto negativo do consumo de álcool sobre a adesão à medicação, não só, em indivíduos infetados com o VIH, como também em indivíduos que sofriam de outras doenças, como por exemplo, a tuberculose (Burman *et al.*, 2007).

Fatores relacionados com a situação clínica

De entre os fatores relacionados com a situação clínica, verificou-se que a situação clínica no momento da entrevista mostrou-se independentemente associada à não-adesão. Em particular, o facto dos doentes apresentarem sintomatologia SIDA revelou-se fator de proteção da não-adesão. Indivíduos com diagnóstico SIDA mostraram-se menos propensos a ser não aderentes. Porém, o estado clínico do doente, isto é, o grau de gravidade da doença, não é um fator consensual na literatura. Existem estudos que mostraram que a adesão à TARc foi significativamente mais baixa entre os indivíduos com menor contagem de células CD4, sugerindo que, como a doença se torna mais intensa, a adesão pode ficar comprometida (Paterson *et al.* (2000), Chesney (2000), Ickovics and Meisler (1997)).

Fatores relacionados com as crenças sobre a doença e os fármacos

No que respeita à percepção da doença por parte do doente, verificou-se que, apesar de não se revelar independentemente associada à não-adesão, a dimensão identidade consta do modelo final de regressão logística múltipla. Para além disso, através do teste de Mann-Whitney, verificou-se associação entre esta dimensão da subescala *IPQ-R* e a adesão (valor- $p=0,017$). O mesmo se verificou para as dimensões controlo de tratamento (valor- $p=0,047$) e representação emocional (0,037). Em particular, verificou-se que, os indivíduos não aderentes possuíam, em média, maior pontuação de cada uma destas três escalas.

Apesar disto, apenas a dimensão controlo de tratamento se verificou independentemente associada à não-adesão (valor- $p=0,001$). Indivíduos que possuíam maior controlo do tratamento, isto é, indivíduos que atribuem importância ao tratamento eram mais propensos a ser não aderentes.

Em relação ao questionário de crenças sobre os medicamentos, verificou-se uma associação entre as preocupações sobre os fármacos e a não-adesão. Após o ajustamento das restantes variáveis, a variável referente às preocupações com os fármacos mostrou-se independentemente associada à não-adesão (valor- $p=0,036$). Quanto maior o nível de preocupação, maior a probabilidade do doente ser não aderente.

Limitações do estudo

Como limitações deste trabalho torna-se perentório referir que o tamanho reduzido da amostra (177 participantes) poderá levar a que não se encontrem todas as associações. O aumento do tamanho da amostra poderia revelar associações que, com o tamanho reduzido da amostra, não se verificaram. Uma segunda limitação consiste no método utilizado para avaliar a adesão. É bastante complexo avaliar a adesão a medicamentos, uma vez que nenhum dos métodos disponíveis se revela infalível. O auto-relato, instrumento utilizado no presente trabalho, tem sido bastante utilizado, porém o ideal é a combinação de metodologias, como por exemplo, o registo de dispensa de medicamentos por parte dos serviços farmacêuticos, sistemas de controlo eletrónico, entre outros. Na avaliação através do auto-relato, a adesão tende a estar sobrestimada, uma vez que os doentes receiam sofrer represálias quando revelam que não cumprem o esquema terapêutico. Todavia, neste estudo, procurou-se minimizar esta limitação com recurso a uma psicóloga treinada e que não fazia da equipa assistencial para a aplicação do questionário.

Uma outra limitação do presente trabalho reside no facto da variável referente à adesão, denominada de AACTG, ter sido dicotomizada em indivíduos aderentes e não aderentes. A escala AACTG define três níveis de adesão: elevada, moderada e fraca. Esta abordagem permite perceber o quão os participantes são aderentes à terapêutica. Tal não é possível

com a dicotomização, uma vez que, de entre os indivíduos considerados não aderentes, encontram-se indivíduos com níveis de adesão distintos, moderados e fracos.

Por último, o desenho do estudo poderá conduzir à sobrestimação da adesão, uma vez que os participantes do estudo foram convidados pelo médico a participar no momento da consulta e, para tal, à partida, tratam-se de indivíduos que comparecem às consultas. Para além disso, os participantes que faltaram na consulta de seguimento foram excluídos do estudo o que implica, igualmente, uma sobrestimação da adesão.

Trabalhos Futuros

Como pesquisa futura seria importante ter acesso ao registo de dispensa de medicação e ao processo clínico dos indivíduos, de modo a ser possível combinar os vários métodos de avaliação da adesão e perceber as concordâncias entre os mesmos. Tal fora referido no capítulo 1, não existe um método infalível, pelo que esta combinação permitiria colmatar as fraquezas de cada um dos métodos através da comparação com os restantes.

Seria, igualmente, relevante, junto da população em estudo, aplicar estratégias que permitissem combater a não-adesão, tendo como foco principal os fatores independentemente associados à não-adesão no presente estudo e, futuramente, avaliar a prevalência de adesão numa amostra selecionada aleatoriamente da população-alvo.

Apêndice A

Referências Bibliográficas

- Agresti, A. (1986) Applying R^2 type measures to ordered categorical data. *Technometrics* **28**, 133–138.
- Amico, K. R., Barta, W., Konkle-Parker, D. J., Fisher, J. D., H., C. D. and Fisher, W. A. (2009) The Brief Illness Perception Questionnaire. *Psychology and Health* **13**(1), 66–75.
- Ammassari, A., Trotta, M. P., Murri, R., Castelli, F., Narciso, P., Noto, P., Vecchiet, J., D’Arminio Monforte, A., Wu, A. W. and Antinori, A. (2002) Corrolates and predictors of adherence to highly active antiretroviral therapy: overview of published literature. *J Acquir Immune Defic Syndr* **3**, 123–127.
- Amorim, M. (2012) *Avaliação do risco oncológico*. Tese de Mestrado, Faculdade de Ciências da Universidade de Lisboa.
- Antunes, C. and Fontaine, A. M. (2005) Perceção de apoio social na adolescência: análise fatorial confirmat?ria da escala Social Support Appraisals. In *Frontiers in Econometrics* **15**(32), 355–366.
- Ash, A. and Shwartz, M. (1999) A useful measure of model performance when predicting a dichotomus outcome. *Statist. Medicine* **18**, 375–384.
- Aykens, J. E., Nease, D. E. J., Nau, D. P., Klinkman, M. S. and Schwenk, T. L. (2005) Adherence to maintenance-phase antidepressant medication as a function of patient beliefs about medication. *Diabetes Care* **3**(1), 23–30.

- Aykens, J. E. and Piette, J. D. (2009) Diabetic patients' medication underuse, illness outcomes, and beliefs about antihyperglycemic and antihypertensive treatments. *Diabetes Care* **32**(1), 19–24.
- Beh, E. J. (2004) Simple Correspondence Analysis: A Bibliographic Review. *International Statistical Review* **72**(2), 257–284.
- Benzécri, J. P. (1973) *L'Analyse des Données*. Paris: Dunod.
- Besch, C. L. (1995) Compliance in clinical trials. *AIDS* **9**, 1–10.
- Blatt, C. R., Citadin, C. B., Souza, F. G., Mello, R. S. and Galato, D. (2009) Avaliação da adesão aos anti-retovirais em um município no Sul do Brasil. *Revista da Sociedade Brasileira de Medicina Tropical* **42**(2), 131–136.
- Bomtempo, N. M. (1999) Estudo dos fatores de risco para o uso irregular do tratamento anti-retroviral, em um serviço público de referência, em Minas Gerais. *Faculdade de Medicina, Universidade Federal de Minas Gerais*.
- Bondelos, J. S., Grothe, K. B., Whitehead, D., Konkle-Parker, D. J., Jones, G. and Brantley, P. J. (2007) Attitudes toward Health Care Providers and Appointment Attendance in HIV/AIDS Patients. *Journal of the association of nurses in AIDS care* **18**(3), 65–73.
- Braithwaite, R. S. and Kendall, J. B. (2010) Influence oh Alcohol Consumption on Adherence to and Toxicity of Antiretroviral Therapy and Survival. *The American Journal of Medicine* **33**(3), 280–287.
- Brito, A. M., Szwarcwald, C. L. and Castilho, E. A. (2006) Fatores associados à interrupção de tratamento anti-retroviral em adultos com AIDS. Rio Grande do Norte, Brasil, 1999-2002. *Revista da Associação Médica Brasileira* **52**(2), 86–92.
- Britten, N. (1994) Patients' ideas about medicines: A qualitative study in a general practice population. *British Journal of General Practice* **44**, 465–468.
- Broadbent, E., Weinman, J., Petrie, K. J. and Main, J. (2006) The Brief Illness Perception Questionnaire. *Psychology and Health* **60**, 631–637.
- Burman, W. J., Cohn, D. L., Rietmeijer, C. A., Judson, F. N., Sbarbaro, J. A. and Reeves, R. R. (2007) Noncompliance with directly observed therapy for tuberculosis. Epidemiology and effect on the outcome of treatment. *Department of Public Health* **111**(5), 1168–1173.

-
- Butler, J. A., Peveler, R. C., Roderick, P., Smith, P. W. F., Horne, R. and Mason, J. C. (2004) Modifiable risk factors for non-adherence to immunosuppressants in renal transplant recipients: A cross sectional study. *Nephrology Dialysis Transplantation* **19**(12), 3144–3149.
- Byrne, M., Walsh, J. and Murphy, A. W. (2005) Secondary prevention of coronary heart disease: Patient beliefs and health-related behaviour. *Journal of Psychosomatic Research* **58**, 5.
- Cabral, S. (2013) *Apontamentos da cadeira de Modelos Estatísticos*.
- Campbell, J. I., Ruano, A. L., Samayoa, B., Estrado Muy, D. L., Arathoon, E. and Young, B. (2010) Adherence to antiretroviral therapy in an Urban, free-care HIV clinic in Guatemala City, Guatemala. *Journal of the International Association of Physicians in AIDS Care* **9**(6), 390–395.
- Campos, L., Guimarães, M. and Remien, R. (2010) Anxiety and depression symptoms as risk factors for non-adherence to antiretroviral therapy in Brazil. *AIDS Behaviour* **14**(2), 289–299.
- Carvalho, C. V., Duarte, D. B., Merchán-Hamann, E., Bicudo, E. and Laguardia, J. (2003) Determinantes da aderência à terapia anti-retroviral combinada em Brasília. *Caderno de Saúde Pública* **19**(2), 593–604.
- Carvalho, C. V., Merchán-Hamann, E. and Matsushita, R. (2007) Determinantes da adesão ao tratamento anti-retroviral em Brasília, DF: um estudo de caso-controle. *Revista da Sociedade Brasileira de Medicina Tropical* **40**(5), 555–565.
- Casel, E. J. (1974) An epidemiological perspective of psychosocial factors in disease etiology. *American Journal of Medicine* **64**(11), 1140–1043.
- Castro, K. G., Ward, J. W., Slutsker, L., Buehler, J. W., Jaffe, H. W. and Berkelman, R. L. (1992) 1993 Revised Classification System for HIV Infection and Expanded Surveillance Case Definition for AIDS Among Adolescents and Adults. *MMWR Recommendations and Reports* **41**, 1–19.
- Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. (1983) *Graphical Methods for Data Analysis*. Wadsworth: .
- Chesney, M. (2000) Factors affecting adherence to Antiretroviral Therapy. *Clinical Infectious Diseases* **30**(2), 171–176.

- Clark, L. and Watson, D. (1991) Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology* **100**(3), 316–336.
- Clemente, S. (2008) *Epidemiologia Molecular da infecção por VIH/SIDA, em Angola*. Lisboa: Faculdade de Medicina.
- Clifford, S., Barber, N. and Horne, R. (165-170) Patient-centred advice is effective in improving adherence to medicines. *Pharmacy World & Science* **28**.
- Clough, L. A., D'Agata, E., Raffanti, S. and Haas, D. W. (1999) Factors that predict incomplete virological response to protease inhibitor-based antiretroviral therapy. *Clin Infect Dis* **29**(1), 75–81.
- Cobb, S. (1976) Social Support as a moderator of life stress. *Psychosomatic Medicine* **38**(5), 300–314.
- Cohen, S. and Mckay, G. (1984) Social Support, Stress, and The Buffering hypothesis; a Theoretical Analysis. *Handbook of Psychology and Health* **4**, 253–268.
- Cohen, S. and Wills, T. (1985) Stress, social support, and the buffering hypothesis. *Psychological Bulletin* **98**(2), 310–357.
- Cohn, S. E., Jiang, H., Mccutchan, J. A., Koletar, S. L., Murphy, R. L., Robertson, K. R., Destmaurice, A. and Currier, J. S. Williams, P. L. (2011) Association of Ongoing Drug and Alcohol Use with Non-Adherence to Antiretroviral Therapy and Higher Risk of AIDS and Death: Results from ACTG 362. *AIDS Care* **23**(6), 775–785.
- Colombrini, M. R., Coleta, M. F. and Lopes, M. H. (2008) Fatores de risco para não adesão ao tratamento com terapia antiretroviral altamente eficaz. *Revista da Escola de Enfermagem da USP* **42**(3), 490–495.
- Colombrini, M. R., Lopes, M. H. and Figueiredo, R. (2006) Adesão à terapia antiretroviral para HIV/AIDS. *Revista da Escola de Enfermagem USP* **40**(4), 576–581.
- Cook, R. L., Sereika, S. M., Hunt, S. C., C., W. W., Erlen, J. A. and Conigliaro, J. (2001) Problem Drinking and Medication Adherence Among Persons with HIV Infection. *J Gen Intern Med* **16**, 83–88.
- Cox, D. R. and Snell, E. J. (1989) *The analysis of Binary Data*. Second edition. London: Chapman and Hall.
- Cox, D. R. and Wermuch, N. (1992) A comment on the coefficient of determination for binary responses. *American Statistician* **46**, 1–4.

- Cramer, J. A., Mattson, R. H., Prevey, M. L., Seller, R. D. and Ouellette, V. L. (1989) How often is medication taken as prescribed? .
- Cutrona, C. E. and Russell, D. W. (1990) *Type of social support and specific stress: Toward a theory of optimal matching*, pp. 319 – 366. New York: Wiley-Interscience.
- Departamento de Doenças Infeciosas: Unidade de referência e Vigilância Epidemiológica. Núcleo de Vigilância Laboratorial de Doenças Infeciosas (Junho de 2013) Infecção VIH/SIDA: a situação em Portugal a 31 de Dezembro de 2012. *Instituto Nacional de Saúde Doutor Ricardo Jorge* (disponível em http://repositorio.insa.pt/bitstream/10400.18/1622/3/Relatorio_SIDA_2012_WEB.pdf) .
- Dias, A. M., Cunha, M., Santos, A., Pinto, A., Silva, A. and Castro, S. (2011) Adesão ao regime terapêutico na doença crónica : revisão da literatura. *Millenium* pp. 201–219.
- Dunbar, M., Ford, G. and Hunt, K. (1998) Why is the receipt social support associated with increased psychosocial distress? An examination of three hypotheses. *Psychology and Health* **13**, 527–544.
- Dunst, C. and Trivette, C. (1990) *Assessment of Social Support in Early Intervention Programs*. New York: Cambridge University Press.
- Eek, F., Merlo, J., Gerdtham, U. and Lithman, T. (2009) Health Care utilisation and attitudes towards health care providers care in subjects reporting environmental annoyance from electricity and chemicals. *Journal of Environment and Public Health* pp. 1–8.
- European Centre for Disease Prevention and Control/WHO Regional Office for Europe (2010) HIV/AIDS surveillance in Europe 2009. *European Centre for Disease Prevention and Control* .
- Fernandes, M. (2011) *Adesão à terapêutica anti-retrovírica em indivíduos seropositivos para o VIH-1: prevalência e fatores associados*. Tese de Mestrado, Faculdade de Medicina da Universidade de Lisboa.
- Ferreira, C. (2013) *Modelos de Regressão: uma aplicação em Medicina Dentária*. Universidade Aberta.
- Figueiras, M., Marcelino, D., Claudino, A., Cortes, M., Maroco, J. and Weinman, J. (2009) Patients' illness schemata of hypertension: The role of beliefs for the choice of treatment. *Psychology & Health* **25**(4), 507–517.

- Figueiras, M. J., Machado, V. A. and Alves, N. C. (2002) Os modelos de senso-comum das cefaleias crônicas nos casais: relação com o ajustamento marital. *Análise Psicológica* (<http://www.scielo.oces.mctes.pt/pdf/aps/v20n1/v20n1a08.pdf>) **1**(XX), 77–90.
- Gifi, A. (1990) *Nonlinear Multivariate Analysis*. New York: Wiley.
- Gomes, J. (2013) *Apontamentos da cadeira de Estatística Aplicada*.
- Gonçalves, M. and Veigas, J. (2009) A influência do exercício físico na ansiedade, depressão e stress. *O Portal dos Psicólogos* (<http://psicologia.com.pt>) .
- Greenacre, M. (1981) *Practical correspondence analysis*. New York: Wiley & Sons.
- Greenacre, M. (1984) *Theory and Application of Correspondence Analysis*. London: Academic Press.
- Greenacre, M. (2005) *From Correspondence Analysis to Multiple and Joint Correspondence Analysis*. Barcelona: Working Paper Number 883, Department of Economics and Business, Universitat Pompeu Fabra.
- Greenacre, M. (2007) *Correspondence Analysis in Practice*. Barcelona, Espanha: Chapman & Hall/CRC.
- Greenacre, M. and Blasius, J. (2006) *Multiple Correspondence Analysis and Related Methods*. Londres: Chapman & Hall/CRC.
- Greenacre, M. J. (1993) *Correspondence Analysis in Practice*. London: Academic Press.
- Greenacre, M. J. and Blasius, J. (1994) *Correspondence Analysis in the Social Sciences*. London: Academic Press.
- Guerreiro, C., Aldir, I., Oliveira, J., Vera, J., Mansinho, K., Marques, L., Mend?o, L., Do-roana, M., Camacho, R., Sarmiento e Castro, R. and Branco, T. (2012) Recomendações Portuguesas para o tratamento da infeção por VIH-1 e VIH-2 2012 .
- Haynes, R. B., Taylor, D. W. and Sackett, D. L. (1979) Compliance in Health Care. *John Hopkins University Press, Baltimore* .
- Hetem, L. A. and Graeff, F. G. (2004) *Transtornos de ansiedade*. Second edition. São Paulo: Atheneu.
- Hill, M. O. (1974) Correspondence Analysis: A Neglected Multivariate Method. *Journal of the Royal Statistical Societes* **23**(3), 340–354.

-
- Horne, R. (1997) *Representations of medication and treatment: Advances in theory and measurement*, pp. 155 – 188. Amsterdam: Harwood.
- Horne, R. (2001a) *Compliance, adherence and concordance*, pp. 166–183. London: Blackwell.
- Horne, R. (2001b) *Non-adherence to medication: Causes and implications for care.*, pp. 111–130. Oxford: Blackwell.
- Horne, R. (2003) *Treatment perceptions and self-regulation.*, pp. 138–154. London: .
- Horne, R. (2006) *Beliefs and adherence to treatment. The challenge for research and practice*, pp. 115–136. Oxford: Oxford University Press.
- Horne, R., Buick, D., Fisher, M., Leake, H., Cooper, V. and Weinman, J. (2004) Doubts about necessity and concerns about adverse effects: Identifying the types of beliefs that are associated with nonadherence to HAART. *International Journal of STD & AIDS* **15**(1), 38–44.
- Horne, R., Frost, S., Hankins, M. and Wright, S. (2001) In the eye of the beholder: Pharmacy students have more positive perceptions of medicines than students of other disciplines. *International Journal of Pharmacy Practice* **9**, 85–89.
- Horne, R. and Weinman, J. (1999) Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. *Journal of Psychosomatic Research* **47**, 555–567.
- Horne, R. and Weinman, J. (2002) Self-regulation and self-management in asthma: Exploring the role of illness perceptions and treatment beliefs in explaining non-adherence to preventive medication. *Psychology and Health* **17**, 17–32.
- Horne, R., Weinman, J. and Hankins, M. (1999) The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology & Health* **14**(1), 1–24.
- Hosmer, D. W. and Lemeshow, S. (2000) *Applied Logistic Regression (2nd Edition)*. New York: John Wiley & Sons.
- Hu, B., Shao, J. and Palta, M. (1999) Pseudo- R^2 in logistic regression model. *Statistica Sinica* **16**, 847–860.
- Ickovics, J. R. and Meisler, A. W. (1997) Adherence in AIDS clinical trials: a framework for clinical research and clinical care. *J Clin Epidemiol* **50**(4), 385–391.

- Infectious Disease Prevention and Control (2012) HIV Trnsmission Risk: A Summary of the Evidence. *Public Health Agency of Canada* .
- Joint United Nations Program on HIV/AIDS (UNAIDS) (2012) World AIDS Day Report 2012. World Health Organization 2012 (disponível em http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/gr2012/jc2434_worldaidsday_results_en.pdf).
- Jordan, R., Gold, L., Cummings, C. and Hyde, C. (2002) Systematic review and meta-analysis of evidence for increasing numbers of drugs in antiretroviral combination therapy. *BMJ* **324(7340): 757**.
- Kessler, R., Price, R. and Wortman, C. (1985) Social Factors in Psychopathology: Stress, Social Support, and Coping Process. *Annual Review of Psychology* **36**, 531–572.
- Kitahata, M. e. a. (2009) Effect of early versus deferred antiretroviral therapy for HIV on survival. *The New England Journal of Medicine* .
- Lau, R. R., Bernard, T. M. and Hartman, K. A. (1989) Further explorations of common sense representations of common illnesses. *Health Psychology* **8**, 195–219.
- Lau, R. R. and Hartman, K. (1983) Common sense representations of common illnesses. *Health Psychology* **2**, 167–185.
- Lazo, M., Gange, S. J., Wilson, T. E., Anastos, K., Ostrow, D. G., Witt, M. D. and Jacobson, L. P. (2007) Patterns and predictors of changes in adherence to highly active antiretroviral therapy: longitudinal study of men and women. *Clin Infect Dis* **45**(10), 1377–1385.
- Leal, I. P., Antunes, R., Pais-Ribeiro, J. and Maroco, J. (2009) Estudo da escala de depressão, ansiedade e stresse para crianças (EADS-C). *Psicologia, Saúde & Doenças* **10**(2), 277–284.
- Lebart, L., Morineau, A. and Warwick, K. M. (1984) *Multivariate Descriptive Statistical Analysis*. Chichester, UK: Wiley.
- Leite, A. (2012) *Análise Longitudinal da mudança terapêutica nos doentes seropositivos para o VIH*. Faculdade de Ciências da Universidade de Lisboa - Tese de Mestrado.
- Leventhal, H. and Diefenbach, M. (1991) The active side of illness cognition. In *Mental Representation in Health and Illness*, eds J. A. Skelton and R. Croyle. New York: Springer-Verlag.

-
- Leventhal, H., Meyer, D. and Nerenz, D. (1980) The common-sense representations of illness danger. pp. 7–30. New York: Pergamon Press.
- Leventhal, H., Nerenz, D. and Strauss, A. (1982) Self-regulation and the mechanisms of symptom appraisal. pp. 55–86. New York: Prodist.
- Leventhal, H., Nerenz, D. R. and Steele, D. J. (1984) Illness representations and coping with health threats. pp. 219–252. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lignani Júnior, L., Greco, D. B. and Carneiro, M. (2001) Avaliação da aderência aos anti-retrovirais em pacientes com infecção pelo HIV/AIDS. *Revista da Sociedade Brasileira de Medicina Tropical* **35**, 495–501.
- Lopes, L. (2011) *Suporte Social e Qualidade de Vida*. Porto: Faculdade de Psicologia e Ciências da Educação.
- Lopes, R., Martins, C., Cruz, P., Mendão, L., Ventura, M., Farinha, H., Campos, M., Rocha, P. and Lucas, R. (2008) *Boas Práticas de Farmácia Hospitalar no âmbito da infecção VIH/SIDA*. Coordenação Nacional para a infecção VIH/SIDA 2008.
- Lovibond, P. and Lovibond, S. (1995) The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy* **33**(3), 335–343.
- Lucas, G. M., Chaisom, R. E. and Moore, R. D. (1999) Highly active antiretroviral therapy in a large urban clinic: risk factors for virologic failure and adverse drug reactions. *Ann Intern Med* **131**(2), 81–87.
- Lucas, G. M., Gebo, K. A., Chaison, R. E. and D., M. R. (2002) Longitudinal assessment of the effects of drug and alcohol abuse in HIV-1 treatment outcomes in an urban clinic. *AIDS* **16**(5), 767–774.
- Maddala, G. S. (1983) Limited-Dependent and Qualitative Variables in Econometrics. *Cambridge University Press*.
- Marôco, J. (2010) *Análise Estatística com o PASW Statistics*. Porto Pinheiro: Report Number.
- Marin-Sanchez, V., Ortega-Valín, L., Pérez-Simón, M. R., Mostaza-Fernández, J. L., Urbina-González, J. J., Rodríguez-María, M., Carro-Fernández, J. A., Cuevas-González, M. J. and Alcoba-Leza, M. (2000) Factores Predictores de no Adherencia to Highly Active Antiretroviral Treatment. *Enfermedades Infecciosas y Microbiología Clínica* **20**(10), 491–497.

- Martins, C. (2005) Impacto do suporte social e dos estilos de coping sobre a percepção subjectiva de bem-estar e qualidade de vida em doentes com l?pus . *O Portal dos Psic?logos* (<http://psicologia.com.pt>) .
- McFadden, D. (1973) Conditional logit analysis of qualitative choice behavior. *In Frontiers in Econometrics* pp. 105–142.
- McKelvey, R. D. and Zavoina, W. (1975) A statistical model for the analysis of ordinal level dependent variables. *Jornal of Mathematical Society* **4**, 103–120.
- Melchior, R., Nemes, M. I. B., Alencar, T. M. D. and Buchalla, C. I. (2007) Desafios da adesão ao tratamento de pessoas vivendo com HIV/AIDS no Brasil. *Revista de Saúde Pública* **41**, 87–93.
- Mitiku, H., Abdosh, T. and Teklemariam, Z. (2013) Factors Affecting Adherence to Anti-retroviral Treatment in Harari National Regional State, Esatern Ethiopia. *ISRN AIDS* .
- Mittlbock, M. and Schemper, M. (1996) Explained Variation of Logisctic Regression Analysis. *Statistics in Medicine* **15**, 1987–1997.
- Moss-Morris, R., Weinman, J., Petrie, K. J., Horne, R., Cameron, L. D. and Buick, D. (2002) The Revised Illness Perception Questionnaire (IPQ-R). *Psychology and Health* (<http://www.uib.no/ipq>) **17**, 1–16.
- Murtagh, F. (2005) *Correspondence Analysis and Data Coding with Java and R*. Boca Raton, FL: Chapman & Hall/CRC.
- Nagelkerke, N. J. D. (1991) A note on a general definition of teh coefficiente of determination. *Biometrika* **78**, 691–693.
- Neame, R. and Hammond, A. (2005) Beliefs about medications: a questionnaire survey of people with rheumatoid arthritis. *Rheumatology* **44**, 762–767.
- Nemes, M. I. B., Souza, M. F. M., Kalichman, A. O., Souza, R. A. and Lopes, J. F. (2000) Avaliação da aderência ao tratamento por anti-retrovirais em usuários de ambulatório do sistema público de assistência à Aids no Estado de São Paulo. *Ministério da Saúde* .
- Nieuwkerk, P. T., Sprangers, M. A., Burger, D. M., Hoetelmans, R. M., Hugem, P. W., Danner, S. A., Van der Ende, M. E., Schneider, M. M., Schrey, G., Meenhorst, P. L., Sprenger, H. G., Kauffmann, R. H., Jambroes, M., Chesney, M. A., De Wolf, F. and lange, J. M. (2000) Limited Patient Adherence to Highly Active Antiretroviral Therapy

-
- for HIV-1 infection in an Observational Cohort Study. *Arch Intern Med* **161**(16), 1962–1968.
- Nishisato, S. (1990) Gleaning in the field of dual scaling. *Psychometrika* **61**, 559–600.
- Norton, W. E., Amico, K. R., Fisher, A. W., Shuper, P. A., Ferer, R. A., Cornamn, D. H., Trayling, C. A., Redding, C. and Fisher, J. (2010) Information-motivation-behavioral skills barriers associated with intencional ARV non-adherence among HIV+ patients in clinical care. *AIDS Care* **22**(8), 979–987.
- Organization, W. H. (2003) *Adherence to long-term therapies*.
- Osborn, C. Y., Amico, K. R., Fisher, W. A., Egede, L. E. and Fisher, J. D. (2010) An Information-Motivation-Behavioral skills analysis od diet and exercise behavior in Puerto Ricans with diabetes. *Journal of Health Psychology* **15**(8), 1201–1213.
- Oxford, C. (1992) Social resources, social support.
- Pais-Ribeiro, J. L. (1999) Escala de Satisfação com o Suporte Social. *Análise Psicológica* **3**, 547–558.
- Pais-Ribeiro, J. L., Honrado, A. and Leal, I. (2004) Contribuição para o estudo da adaptação portuguesa das escalas de ansiedade, depressão e stress (EADS) de 21 itens de Lovibond e Lovibond. *Psicologia, Sa?de & Doen?as* **5**(2), 229–239.
- Palella, F., Jr., M. and Delaney, K. (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *The New England Journal of Medicine* pp. 853–860.
- Paterson, D. L., Swindells, S., Mohr, J., Breste, M., Vergis, E. N., Squier, C., Wagener, M. M. and Singh, N. (2000) Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Ann Intern Med* **133**(1), 21–30.
- Paula, C. C., Brum, C. N., Zuge, S. S., Rodrigues, A. P., Tolentino, L. C. and Padoin, S. M. (2012) Caracterização da morbimortalidade de crianças com HIV/AIDS em serviço de referência no Sul do Brasil. *Revista Saúde (Santa Maria)* **38**(2), 25–36.
- Pedroso, R. S. and Sbardelloto, G. (2008) Qualidade de vida e suporte social em pacientes renais crônicos: revisão teórica. *Psicópio: Revista Virtual de Psicologia Hospitalar e da Saúde* **7**(32), 32–35.
- Peltzer, K., Fien-du Preez, N., Ramlagan, S. and Anderson, J. (2010) Antiretroviral treatment adherence among HIV patients in KwaZulu-Natal, South Africa. *BMC Public Health* **10**.

- Pereira, L. B., Albuquerque, J. R., Santos, J. M., Lima, F. L. and Saldanha, A. A. (2012) Fatores Sociodemográficos e Clínicos associados à TARV e à contagem T-CD4. *Revista Brasileira de Ciências da Saúde* **16**(2), 149–160.
- Pereira, M. G., Pedras, S. and Machado, J. (2013) Validação do Questionário Crenças acerca da Medicação em Pacientes Diabéticos Tipo 2. *Psicologia: Teoria e Pesquisa* **29**(2), 229–236.
- Petrie, K. J., Broadbent, E. and Kydd, R. (2008) Illness perceptions in mental health: issues and potencial application. *Journal of Mental Health* **17**(6), 559–564.
- Phatak, H. M. and Thomas, J. (2006) Relationships between beliefs about medications and nonadherence to prescribed chronic medications. *The Annals of Pharmacotherapy* **40**, 1737–1742.
- Pietrukowick, M. (2001) *Apoio social e religião: uma forma de enfrentamento dos problemas de saúde*. Escola Nacional de Saúde Pública - Tese de Mestrado.
- Pomerantz, R. J. and Horn, D. L. (2013) Twenty years of therapy for HIV-1 infection. *Nature Medicine* **9**(7), 867–873.
- Powell, T. J. and Enright, S. J. (1990) *Anxiety and stress management*. Routledge (London and New York).
- Reis, A. C., Lencastre, L., Guerra, M. P. and Remor, E. (2009) Adaptação portuguesa do questionário para a avaliação da adesão ao tratamento anti-retrovírico - VIH (CEAT-VIH). *Psicologia, Saúde & Doenças* **10**(2), 175–191.
- Reis, J. C. (2005) O que é a saúde? Significações pessoais, modelos científicos e educação para a saúde. *Análise Psicológica* **23**(2).
- Remor, E., Milner-Moskovics, J. L. and Preussler, G. (2007) Adaptação brasileira do "Cuestionario para la Evaluaci?n de la Adhesion al Tratamiento Antiretroviral. *Revista de Saúde Pública* **41**(5), 685–694.
- Riera, M., De la Fuente, L., Puigventos, F., Villalonga, C., Salas, A., Pareja, A. and De la Pe?a, A. (eds) (1999) *Accuracy of different indirect methods to detect anti-retroviral adherence needed to obtain a good viral control at 9 months of follow-up*, Lisbon. Seventh European Conference on Clinical Aspects and Treatment of HIV-infection, Abstract Book, European AIDS Clinical Society.

- Rigotto, D. (2006) *Evidências de validade entre suporte familiar, suporte social e autoconceito*. Psicologia da Universidade de São Francisco - Dissertação de Mestrado (Programa de pós graduação *Stricto Sensu*).
- Romeu, G. A., Tavares, M. M., Carmo, C. P., Magalhães, K. N., Nobre, A. C. and Matos, V. C. (2012) Avaliação da adesão à terapia antiretroviral de pacientes portadores de HIV. *Rev. Bras. Farm. Hosp. Serv. Saúde* **3**(1), 37–41.
- Ross, S., Walker, A. and MacLeod, M. J. (2004) Patient compliance in hypertension: The role of illness perceptions and treatment beliefs. *Journal of Human Hypertension* **18**, 607–613.
- Sabaté, E. (2001) Long-term Therapies: Policy for action. *World Health Organization* .
- Saldanha, J. S., Andrade, C. S. and Beck, S. T. (2009) Grau de adesão ao tratamento com anti-retrovirais entre indivíduos HIV positivos atendidos no hospital universitário de Santa Maria. *Revista Saúde (Santa Maria)* **35**(1), 4–9.
- Samet, J. H., Horton, N. J., Meli, S., Freedberg, K. A. and Palepu, A. (2004) Alcohol consumption and antiretroviral adherence among HIV-infected persons with alcohol problems. *Alcoholism: Clinical & Experimental Research* **28**(4), 572–577.
- Santos, C. and Pais-Ribeiro, J. and Lopes, C. (2003) Adaptação e validação do "Revised Illness Perception Questionnaire" em doentes oncológicos. *Arquivos de Medicina* **17**(4), 136–147.
- Santos, P. (2008) Mudanças de tratamento: Quando o tratamento anti-retroviral falha: terapêutica de segunda e terceira linha e resistência aos medicamentos. *Grupo Português de Ativistas sobre tratamentos de VIH/SIDA (GAP)* .
- Santos, P. (2013) Introdução à terapêutica de combinação para a infeção pelo VIH. *Grupo Português de Ativistas sobre tratamentos de VIH/SIDA (GAP)* .
- Sarason, I. G., Levine, H. M., Basham, R. B. and Sarason, B. R. (1983) Assessing Social Support: The Social Support Questionnaire. *Journal of Personality and Social Psychology* **44**(1), 127–139.
- Schoeber, R. and Lacroix, J. M. (1991) *Lay illness models in the enlightenment and the 20th Century: some historical lessons*. New York: Springer-Verlag.
- Seidl, E., Rossi, W., Viana, K., Meneses, A. K. and Meireles, E. (2005) Crianças e adolescentes vivendo com HIV/Aids e suas famílias: aspectos psicossociais e enfrentamento. *In Frontiers in Econometrics* **21**(3), 279–288.

- Seidl, E. M., Melch?ades, A., Farias, V. and Brito, A. (2007) Pessoas vivendo com HIV/AIDS: Variáveis associadas à adesão ao tratamento anti-retroviral. *Caderno de Saúde Pública* **23**(10), 2305–2316.
- Serra, A. (1999) *O stress na vida de todos os dias*. Coimbra: Edições do Autor.
- Shatland, E. S., Kleinman, K. and Cain, E. M. (2002) *Onde More Time About R^2 Measures of Fit in Logistic Regression*. Buffalo, New York: .
- Singer, J. E. and Lord, D. (1984) The Role of Social Support in Coping With Chronic or Life-Threatning Illness. *Handbook of Psychology and Health* **4**, 269–278.
- Siqueira, M. M. (2008) Construção e validação da escala de percepção de suporte social. *Psicologia em estudo* **13**(2), 381–388.
- Sousa, M. R. M. G. C. (2003) *Estudo dos conhecimentos e representações de doença associados à adesão à terapêutica nos diabéticos tipo 2*. Dissertação de Mestrado, Universidade do Minho (Disponível em <http://repositorium.sdum.uminho.pt/handle/1822/668>).
- Symister, P. and Friend, R. (2003) The influence of social support and problematic support on optimism and depression in chronic illness: a prospective study evaluating self-esteem as a mediator. *Health Psychology* **22**(2), 123–129.
- Talam, N. C., Gatongi, P., Rotich, J. and kimaiyo, S. (2008) Factors affecting antiretroviral drug adherence among HIV/AIDS adult patients attending HIV/AIDS clinic at Moi Teaching and Referral Hospital, Eldoret, Kenya. *East African journal of public health* **5**(2), 74–78.
- Teixeira, J. A. and Leal, I. (1993) Psicologia da saúde: Contexto e intervenção. *Análise Psicológica* **8**(4), 453–458.
- Treharne, G. J., Lyons, A. C. and Kitas, G. (2004) Medication adherence in rheumatoid arthritis, effects of psychological factors. *Psychology of Health Medicine* **9**, 337–349.
- Tucker, J. S., Burnam, M. A., Sherbourne, C. D., Kung, F. Y. and Gifford, A. L. (2003) Substance use and mental health correlates of nonadherence to antiretroviral medications in a sample of patients with human immunodeficiency virus infection. *The American Journal of Medicine* **114**(7), 573–580.
- Turkman, M. A. and Silva, G. (2000) *Modelos Lineares Generalizados - da teoria à prática*. Lisboa: Edições SPE.

- Urquhart, J. (1991) Compliance and clinical trials. *The Lancet* **337**, 1224–1225.
- Vale, M. S. (1999) *Avaliação da Aderência ao Tratamento de Pacientes com HIV/AIDS no Centro de Saúde de Taguatinga Nº 4*. Brasília: Monografia de Especialização, Universidade de Brasília.
- Van Servellen, G., Chang, B. and Garcia, L. Lombardi, E. (2002) Individual and system level factors associated with treatment nonadherence in human immunodeficiency virusinfected men and women. *AIDS Patient Care STDS* **16**, 269–281.
- Vaux, A. (1988) *Social Support: Theory, research and intervention*. New York: Praeger.
- Vermeire, E., Hearnshaw, H., Von Royen, P. and Denekens, J. (2001) Patient adherence to treatment: Three decades of research. A comprehensive review. *Journal of Clinical Pharmacy and Therapeutics* **26**, 331–342.
- Watson, D., Weber, K., Assenheimer, J., Clark, L., Staruss, M. and McCormick, R. (1995) Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology* **104**(1), 3–14.
- Weinman, J., Petrie, K. J., Moss-Morris, R. and Horne, R. (1996) The Illness Perception Questionnaire: A new method for assessing the cognitive representation of illness. *Psychology and Health* (<http://www.uib.no/ipq/>) **11**, 431–435.
- Wethington, E. and Kessler, R. C. (1986) Perceived support, received support, and adjustment to stressful life events. *Journal of Health and Social Behavior* **27**, 78–89.
- Yach, D. (2003) Adherence to long-term therapies: Evidence for action. *World Health Organization* .
- Zheng, B. Y. and Agresti, A. (1999) Summarizing the predictive power of a generalized linear model. *Statist. Medicine* **19**, 1771–1781.
- Zweig, M. H. and Campbell, G. (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* **39**(4), 561–577.

Apêndice B

Anexos

Tabela B.1: Grelha de correção da Escala de Satisfação com o Suporte Social (ESSS)

	Discordo plenamente	Discordo	Não concordo nem discordo	Concordo	Concordo plenamente
1 - Por vezes sinto-me só no mundo e sem apoio. (IN)	5	4	3	2	1
2 - Não saio com amigos tantas vezes quanto gostaria. (AS)	5	4	3	2	1
3 - Os amigos não me procuram tantas vezes quanto eu gostaria. (AS)	5	4	3	2	1
4 - Quando penso em desabafar com alguém, encontro facilmente amigos com quem o fazer. (IN)	1	2	3	4	5
5 - Mesmo nas situações mais embaraçocas, se precisar de apoio de emergência tenho várias pessoas a quem recorrer. (IN)	1	2	3	4	5
6 - Às vezes sinto falta de alguém verdadeiramente íntimo que me compreenda e com quem possa desabafar coisas íntimas. (IN)	5	4	3	2	1
7 - Sinto falta de atividades sociais que me satisfaçam	5	4	3	2	1
8 - Gostava de participar mais em atividades de organizações (p. ex. clubes desportivos, partidos políticos, etc). (AS)	5	4	3	2	1
9 - Estou satisfeito com a forma como me relaciono com a minha família. (SF)	1	2	3	4	5
10 - Estou satisfeito com a quantidade de tempo que passo com a minha família. (SF)	1	2	3	4	5
11 - Estou satisfeito com o que faço em conjunto com a minha família. (SF)	1	2	3	4	5
12 - Estou satisfeito com a quantidade de amigos que tenho. (SA)	1	2	3	4	5
13 - Estou satisfeito com a quantidade de tempo que passo com os meus amigos. (SA)	1	2	3	4	5
14 - Estou satisfeito com as atividades e coisas que faço com o meu grupo de amigos. (SA)	1	2	3	4	5
15 - Estou satisfeito com o tipo de amigos que tenho. (SA)	1	2	3	4	5

Tabela B.2: Grelha de correção da Escala de Ansiedade, Depressão e Stress (EADS)

	Não se aplicou a mim	Aplicou-se a mim algumas vezes	Aplicou-se a mim muitas vezes	Aplicou-se a mim a maior parte das vezes
1 - Tive dificuldade em me acalmar. (S)	0	1	2	3
2 - Senti a minha boca seca. (A)	0	1	2	3
3 - Não consegui sentir nenhum sentimento positivo. (D)	0	1	2	3
4 - Senti dificuldade em respirar. (A)	0	1	2	3
5 - Tive dificuldade em tomar iniciativa para fazer as coisas. (D)	0	1	2	3
6 - Tive tendência a reagir em demasia em determinadas situações. (S)	0	1	2	3
7 - Senti tremores (p.ex., nas mãos). (A)	0	1	2	3
8 - Senti que estava a utilizar muita energia nervosa. (S)	0	1	2	3
9 - Preocupei-me com situações em que podia entrar em pânico e fazer figura ridícula. (A)	0	1	2	3
10 - Senti que não tinha nada a esperar do futuro. (D)	0	1	2	3
11 - Dei por mim a ficar agitado. (S)	0	1	2	3
12 - Senti dificuldade em me relaxar. (S)	0	1	2	3
13 - Senti-me desanimado e melancólico. (D)	0	1	2	3
14 - Estive intolerante em relação a qualquer coisa que me impedisse de terminar aquilo que estava a fazer (S).	0	1	2	3
15 - Senti-me quase a entrar em pânico. (A)	0	1	2	3
16 - Não fui capaz de ter entusiasmo para nada. (D)	0	1	2	3
17 - Senti que não tinha muito valor como pessoa. (D)	0	1	2	3
18 - Senti que por vezes estava sensível. (S)	0	1	2	3
19 - Senti alterações no meu coração sem fazer exercício físico. (A)	0	1	2	3
Senti-me assustado sem ter tido uma boa razão para isso. (A)	0	1	2	3
Senti que a vida não tinha sentido. (D)	0	1	2	3

Tabela B.3: Grelha de correção da subescala identidade da escala IPQ-R

	Desde o início da sua doença, já teve este sintoma?		Se sim, acha que ele está relacionado com a sua doença?	
	Sim	Não	Sim	Não
Dores	-	-	1	0
Dores de garganta	-	-	1	0
Náuseas	-	-	1	0
Falta de ar	-	-	1	0
Pesa de peso	-	-	1	0
Fadiga (cansaço)	-	-	1	0
Rigidez nas articulações	-	-	1	0
Olhos inflamados	-	-	1	0
Dificuldade em respirar	-	-	1	0
Dores de cabeça	-	-	1	0
Indisposição de estômago	-	-	1	0
Dificuldade em dormir	-	-	1	0
Tonturas	-	-	1	0
Perda de forças	-	-	1	0

Tabela B.4: Grelha de correção das restantes subescalas da escala IPQ-R

	Discordo plenamente	Discordo	Não concordo nem discordo	Concordo	Concordo plenamente
A minha doença dura pouco tempo.	5	4	3	2	1
A minha doença parece ser mais permanente que temporária.	1	2	3	4	5
A minha doença vai durar muito tempo.	1	2	3	4	5
Esta doença vai passar depressa.	5	4	3	2	1
Acho que a minha doença vai durar para o resto da minha vida.	1	2	3	4	5
A minha doença é grave.	1	2	3	4	5
A minha doença tem grandes consequências para a minha vida.	1	2	3	4	5
A minha doença não tem grande efeito na minha vida.	5	3	3	2	1
A minha doença afeta seriamente a forma como os outros me vêem	1	2	3	4	5
A minha doença tem sérias consequências económicas.	1	2	3	4	5
A minha doença causa dificuldades aqueles que me são próximos.	1	2	3	4	5
Há muita coisa que eu posso fazer para controlar os meus sintomas.	1	2	3	4	5
O que eu faço pode determinar se a minha doença melhora ou piora.	1	2	3	4	5
A evolução da minha doença depende de mim.	1	2	3	4	5
Nada do que eu faça irá afetar a minha doença.	5	4	3	2	1
Eu tenho o poder de influenciar a minha vida.	1	2	3	4	5
As minhas ações não terão qualquer efeito no resultado da minha doença	5	4	3	2	1
A minha doença vai melhorar com o tempo.	5	4	3	2	1
Há pouco que se possa fazer para melhorar a minha doença.	5	4	3	2	1
O tratamento é eficaz na cura da minha doença.	1	2	3	4	5
Os efeitos negativos da minha doença poderão ser prevenidos ou evitados pelo tratamento.	1	2	3	4	5
O tratamento pode controlar a minha doença.	1	2	3	4	5
Não há nada que possa ajudar a minha situação.	5	4	3	2	1
Os sintomas da minha doença confundem-me.	1	2	3	4	5
A minha doença é um mistério para mim.	1	2	3	4	5
Não compreendo a minha doença.	1	2	3	4	5
A minha doença não faz qualquer sentido para mim.	1	2	3	4	5
Tenho uma ideia clara sobre a minha doença.	5	4	3	2	1
Os sintomas da minha doença mudam bastantes de dia para dia.	1	2	3	4	5
Os sintomas da minha doença vêm e vão em ciclos	1	2	3	4	5
A minha doença é muito imprevisível.	1	2	3	4	5
A minha doença passa por fases em que melhora ou piora.	1	2	3	4	5
Fico deprimido quando penso sobre a minha doença.	1	2	3	4	5
Quando penso sobre a minha doença fico perturbado.	1	2	3	4	5
A minha doença faz-me sentir zangado.	1	2	3	4	5
A minha doença não me preocupa.	5	4	3	2	1
Ter esta doença torna-me ansioso.	1	2	3	4	5
A minha doença faz-me sentir medo.	1	2	3	4	5

Tabela B.5: Grelha de correção da escala de crenças acerca dos fármacos

	Discordo plenamente	Discordo	Não concordo nem discordo	Concordo	Concordo plenamente
Atualmente, a minha saúde depende da medicação anti-retrovírica que me foi prescrita. (N)	1	2	3	4	5
Preocupa-me ter que tomar esta medicação. (P)	1	2	3	4	5
A minha vida seria impossível sem esta medicação. (N)	1	2	3	4	5
Sem esta medicação eu ficaria muito doente. (N)	1	2	3	4	5
Por vezes, preocupo-me com o efeitos a longo prazo. (P)	1	2	3	4	5
A medicação é um mistério para mim. (P)	1	2	3	4	5
No futuro, a minha saúde vai depender deste medicação. (N)	1	2	3	4	5
Esta medicação perturba a minha vida. (P)	1	2	3	4	5
Por vezes, preocupa-me ficar demasiado dependente desta medicação. (P)	1	2	3	4	5
Esta medicação protege-me de ficar pior. (N)	1	2	3	4	5
Esta medicação provoca-me efeitos secundários desagradáveis. (P)	1	2	3	4	5
Usar esta medicação é embaraçoso. (P)	1	2	3	4	5
Não tomar esta medicação durante um dia não tem importância a longo prazo. (N)	5	4	3	2	1
Esta medicação aumenta a esperança no futuro. (N)	1	2	3	4	5
Não é provável que eu tenha um efeito secundário desagradável causado por esta medicação durante o próximo mês. (P)	5	4	3	2	1
Tomar esta medicação está a ser muito pior do que esperava. (P)	1	2	3	4	5
Recebi informação suficiente sobre a medicação anti-retrovírica. (P)	5	4	3	2	1
Esta medicação mantém-me vivo. (N)	1	2	3	4	5
O gosto da medicação anti-retrovírica faz-me sentir mal. (P)	1	2	3	4	5

Data:

DD	MM	AAAA

ID Participante

--	--	--	--	--	--	--	--	--	--

QUESTIONÁRIO INICIAL

Este questionário está incluído no **estudo sobre a utilização medicação anti-retrovirica**, no qual aceitou participar.

O questionário tem como objectivo ajudar-nos a compreender como é que as pessoas utilizam estes medicamentos, quais as dificuldades que sentem com a medicação e qual a opinião que têm sobre a sua doença.

Não existem respostas certas ou erradas e a sua participação é confidencial.

Se tiver alguma dúvida, chame uma entrevistadora.

Por favor faça os possíveis para responder a todas as perguntas.

Instruções de Preenchimento	Para efeito de tratamento dos dados (leitura óptica), caneta ou esferográfica preta ou azul.	Para responder, No caso de pretender corrigir

Não escrever nas caixas sombreadas (caixas cinzentas, exemplo:

--	--

No final, a entrevistadora vai rever consigo eventuais dúvidas que tenha tido no preenchimento do questionário.

Desde já, muito obrigado pela sua colaboração!

Bloco A

Para começar, pedimos-lhe que nos indique alguma informação geral para caracterização sócio-demográfica

1. Qual é a sua data de nascimento?

DD	MM	AAAA

2. Sexo?

- ☐ Masculino
☐ Feminino

3. Qual é a sua Nacionalidade?

--	--

4. Em que País nasceu?

--	--

5. Com qual etnia/raça se identifica mais?

- ☐ Branco de origem europeia
☐ Branco de origem não-europeia
☐ Negro
☐ Mulato
☐ Romani/Cigano
☐ Asiático, Indiano
☐ Asiático, Chinês/Japoneses
☐ Outra

--	--

6. Qual é o seu estado civil?

- ☐ Solteiro
☐ Casado / União de Facto/vive maritalmente há pelo menos 2anos
☐ Divorciado
☐ Viúvo

7. Tem filhos menores a seu cargo?

- ☐ Sim
☐ Não

8. Mora sozinho?

- ☐ Sim
☐ Não → 8.a.Com quantas pessoas mora (contando consigo)?

--	--

8.b. Indique as pessoas com quem mora

- ☐ Marido/Esposa ☐ Filho/Filha → Quantos?

--	--

☐ Mãe ☐ Irmão/Irmã
☐ Pai ☐ Outra pessoa de família
☐ Sogro/sogra ☐ Outra pessoa. Quem?

--	--

☐ Avó/Avô

Estudo sobre utilização de medicação anti-retrovirica

12. As pessoas têm diferentes estilos de vida. As questões que se seguem referem-se à sua actividade física e ao seu consumo de álcool, tabaco e outros, nos últimos 30 dias:

12.a. Com que frequência faz exercício físico? (caminhada, natação, jogging, ciclismo, futebol, ténis, etc.)

- ☐ Diariamente
☐ Mais de 4 vezes por semana
☐ 3 a 4 vezes por semana
☐ 1 a 2 vezes por semana
☐ 2 a 3 vezes por mês
☐ 1 vez por mês
☐ Nunca

☐ O médico aconselhou-me a não fazer exercício físico

12.b. Durante os últimos 30 dias e de uma forma geral, com que frequência ingeriu um bebida com álcool? Por bebida referimo-nos a uma lata ou copo de cerveja, um copo de vinho, digestivo ou bebida branca.

- ☐ Diariamente
☐ Mais de 4 vezes por semana
☐ 3 a 4 vezes por semana
☐ 1 a 2 vezes por semana
☐ 2 a 3 vezes por mês
☐ 1 vez por mês
☐ Nunca -> Passar para a questão 13

12.c. Tendo em conta as ocasiões em que ingeriu bebidas com álcool, quantas bebidas em média consumiu numa mesma ocasião?

- ☐ 1 ou 2 bebidas por ocasião
☐ 3 ou 4 bebidas por ocasião
☐ 5 ou 6 bebidas por ocasião
☐ 7 ou 8 bebidas por ocasião
☐ 9 a 11 bebidas por ocasião
☐ 12 ou mais bebidas por ocasião

12.d. Durante os últimos 30 dias, com que frequência consumiu 5 ou mais bebidas alcoólicas, num período até 4 horas?

- ☐ Diariamente
☐ Mais de 4 vezes por semana
☐ 3 a 4 vezes por semana
☐ 1 a 2 vezes por semana
☐ 2 a 3 vezes por mês
☐ 1 vez por mês
☐ Nunca

13. Como caracterizar os seus hábitos tabágicos?

☐ Nunca fumou

Há quanto tempo deixou de fumar? anos
☐ Deixou de fumar - (Se preferir indique o ano)

☐ Fuma ocasionalmente -> Quantos cigarros por semana?

cigarros/semana

☐ Fuma -> Quantos cigarros por dia?

cigarros por dia

13.b. Por favor indique "Sim" ou "Não" para cada pergunta, tendo em conta os últimos 6 meses.

	Sim	Não
Consumiu maquiagem durante os últimos 6 meses?	<input type="checkbox"/>	<input type="checkbox"/>
Consumiu anfetaminas (speeds) durante os últimos 6 meses?	<input type="checkbox"/>	<input type="checkbox"/>
Consumiu cocaína durante os últimos 6 meses?	<input type="checkbox"/>	<input type="checkbox"/>
Consumiu heroína durante os últimos 6 meses?	<input type="checkbox"/>	<input type="checkbox"/>

14. Por favor marque "Sim" ou "Não"

	Sim	Não
Alguma vez esteve em tratamento com meladona?	<input type="checkbox"/>	<input type="checkbox"/>
Se sim, está actualmente a fazer tratamento?	<input type="checkbox"/>	<input type="checkbox"/>

3 / 12

49005



Estudo sobre utilização de medicação anti-retrovirica

Responda às seguintes perguntas colocando um "X" no quadrado que achar mais apropriado. Por favor escolha apenas uma opção para cada pergunta

Bloco B

As pessoas têm muitas vezes dificuldade em tomar a sua medicação de acordo com as indicações do médico. Por exemplo, algumas pessoas não conseguem tomar todos os comprimidos receitados no número de vezes indicadas.

1. Há quanto tempo está a tomar medicação anti-retrovirica?

- ☐ Esta foi a primeira consulta que me indicaram este tipo de medicação
☐ Há 3 meses ou menos
☐ Há 6 meses ou menos
☐ Há 12 meses ou menos
☐ Há mais de 12 meses

2. Nas afirmações que são apresentadas de seguida, qual a que melhor o descreve actualmente, em relação à toma da medicação anti-retrovirica?

- ☐ Eu não estou a tomar medicação e não estou a pensar começar a tomá-la como indicado
☐ Eu não estou a tomar medicação mas estou a pensar começar a tomá-la como indicado
☐ Eu não estou a tomar medicação mas vou começar a tomá-la como indicado
☐ Eu estou a tomar medicação como indicado.

3. Tendo em conta a **medicação anti-retrovirica** (para a infecção VIH/SIDA) até que ponto tem a certeza de que...

	Não tem qualquer certeza	Tem alguma certeza	Tem certeza	Tem a certeza absoluta
... a medicação terá um efeito positivo na sua saúde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... o VIH toma-se resistente caso não tome a medicação exactamente como prescrita?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... vai ser capaz de tomar toda a medicação tal como foi indicada pelo médico?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Se vai iniciar a toma de medicação nesta consulta, passe para o Bloco C (na página 8)

4. Indique, por favor o nome dos medicamentos anti-retroviricos que está a tomar neste momento. Se não se lembrar do nome de algum medicamento, escreva "Não sei" na linha correspondente.

1- 3-
 2- 4-

☐ Não me lembro do nome de nenhum medicamento

5. Existe algum medicamento que seja mais difícil tomar?

☐ Sim ☐ Não -> Passe para a Pergunta 6

5.b. Se sim, qual(uais) o(s) nome(s)?

5c. Porque motivo é difícil tomar esse medicamento?

- ☐ Sabor
☐ Horário
☐ Receto de efeitos adversos
☐ Incompatibilidade com o consumo de alimentos
☐ Incompatibilidade com o consumo de álcool/drogas
☐ Outro motivo. Qual?

4 / 12

49005



6. Como avalia a quantidade de informação que tem sobre a medicação?

Nada/Nenhum(a)	Pouca(a)	Algum(a)	Bastante	Muito(a)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7.

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. Tem alguma estratégia para se lembrar de tomar a medicação?

☐ Sim --> Qual?

☐ Não

9. Quando foi a última vez em que falhou alguma toma da medicação?

☐ Durante a última semana

☐ Há 1 ou 2 semanas atrás

☐ Há 2 ou 4 semanas atrás

☐ Há 1 ou 3 meses atrás

☐ Há mais de 3 meses

☐ Nunca deixou de tomar a medicação

10. As pessoas podem deixar de tomar a medicação por várias razões. Segue-se uma lista de possíveis razões pelas quais isso lhe possa ter acontecido.

Tendo em conta os **últimos 30 dias**, com que frequência não tomou a sua medicação...

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

11. Considere agora os **últimos 7 dias**, contando com o dia de ontem.

Em alguma ocasião...

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Estudo sobre utilização de medicação anti-retrovírica

Bloco C

Obrigada pela informação que indicou até agora.

Gostáramos agora de conhecer a sua opinião sobre a sua doença

1. A pergunta seguinte refere-se à forma como pensa ter sido infectado(a) pelo VIH. Assinale com um X a opção que melhor se aplica ao seu caso. Pode assinalar mais que uma opção, se necessário.

☐ Relações sexuais com um homem que era VIH+

☐ Relações sexuais com uma mulher que era VIH+

☐ Partilha de seringas com uma pessoa que era VIH

☐ Transfusão sanguínea ou outro procedimento médico

☐ Outro (acidente no trabalho, etc.) Qual? (especifique)

A seguir, apresentamos uma lista de sintomas que poderá ter tido ou não desde o início da sua doença.

2. Por favor indique (sim ou não) se já teve algum destes sintomas desde o início da sua doença, e se acha que estes sintomas estão relacionados com ela.

	Desde o início da sua doença, já teve este sintoma?		Se sim, já teve este sintoma, acha que ele está relacionado com sua doença?	
Dores	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Dores de garganta	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Náuseas	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Falta de ar	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Perda de peso	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Fadiga (cansaço)	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Rigidez das articulações	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Olhos inflamados	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Dificuldade em respirar	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Dores de cabeça	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Indisposição de estômago	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Dificuldade em dormir	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Tonturas	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma
Perda de forças	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Não <input type="checkbox"/> Não tive este sintoma

Estudo sobre utilização de medicação anti-retrovírica

12. Algumas pessoas esquecem-se de tomar a sua medicação durante os fins-de-semana. No seu caso, falhou alguma da sua medicação anti-retrovírica no último fim-de-semana (último sábado ou domingo)?

☐ Sim

☐ Não

13. Considerando agora os últimos 4 dias, indique quantos dias não tomou a sua medicação?

Caso tenha tomado apenas uma parte da medicação (por exemplo: apenas 1 comprimido, quando teria de tomar 3 comprimidos) em um ou mais dias, por favor considere esses dias como não tendo tomado a sua medicação.

☐ Nenhum dia ☐ 1 dia ☐ 2 dias ☐ 3 dias ☐ 4 dias

13.a. A maior parte da medicação anti-retrovírica tem que ser tomada segundo um esquema, por ex. "2 vezes por dia" ou "3 vezes por dia" ou "de 8 em 8 horas". Considerando os últimos 4 dias, com que frequência cumpriu o esquema específico da sua medicação?

☐ Sim

☐ Não

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

13.b. Alguma da sua medicação anti-retrovírica tem instruções especiais, tais como "tomar com comida" ou "com o estômago vazio" ou "com muitos líquidos"?

- ☐ Sim
- ☐ Não

13.c. Se sim, com que frequência seguiu essas instruções especiais nos últimos 4 dias?

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

14. Em geral, como avalia a sua satisfação com a medicação desde que a começou a tomar?

Muito Insatisfeito	Insatisfeito	Nem Satisfeito, Nem Insatisfeito	Satisfeito	Muito Satisfeito
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

15. Que avaliação faz de si mesmo relativamente à toma da medicação?

Nada Cumpridor	Pouco Cumpridor	Moderadamente Cumpridor	Bastante Cumpridor	Muito Cumpridor
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

16. Até que ponto está satisfeito com a relação que tem com o médico que o acompanha na infeção VIH/SIDA?

Muito Insatisfeito	Insatisfeito	Nem Satisfeito, Nem Insatisfeito	Satisfeito	Muito Satisfeito
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. De seguida, apresentamos um conjunto de afirmações sobre a forma com as pessoas vêem a sua própria doença. Pedimos-lhe que indique em que medida concorda ou discorda com cada afirmação, colocando um X no quadrado que ache apropriado.

	Discorda Plenamente	Discorda	Não Concorda Nem Discorda	Concorda	Concorda Plenamente
A minha doença dura pouco tempo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença parece ser mais permanente que temporária	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença vai durar muito tempo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esta doença vai passar depressa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acho que a minha doença vai durar para o resto da minha vida	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença é grave	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença tem grandes consequências para a minha vida	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença não tem grande efeito na minha vida	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença afecta seriamente a forma como os outros me veem	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença tem sérias consequências económicas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença causa dificuldades aqueles que me são próximos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Há muita coisa que eu posso fazer para controlar os meus os sintomas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O que eu faço pode determinar se a minha doença melhora ou piora	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A evolução da minha doença depende de mim	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nada do que eu faça irá afectar a minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu tenho o poder de influenciar a minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
As minhas acções não terão qualquer efeito no resultado da minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença vai melhorar com o tempo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Há pouco que se possa fazer para melhorar a minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O tratamento é eficaz na cura da minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Os efeitos negativos da minha doença poderão ser prevenidos ou evitados pelo tratamento	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O tratamento pode controlar a minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



	Discorda Plenamente	Discorda	Não Concorda Nem Discorda	Concorda	Concorda Plenamente
Não há nada que possa ajudar a minha situação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Os sintomas da minha doença confundem-me	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença é um mistério para mim	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Não compreendo a minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença não faz qualquer sentido para mim	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tenho uma ideia clara sobre a minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Os sintomas da minha doença mudam bastante de dia para dia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Os sintomas da minha doença vêm e vão em ciclos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença é muito imprevisível	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença passa por fases em que melhora ou piora	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fico deprimido(a) quando penso sobre a minha doença	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Quando penso sobre a minha doença fico perturbado(a)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença faz-me sentir zangado(a)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença não me preocupa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ter esta doença torna-me ansioso(a)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha doença faz-me sentir medo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bloco D

As questões seguintes referem-se ao sente sobre as suas **relações interpessoais e ao apoio** para a toma da medicação

1. Com que frequência os seus amigos ou familiares o ajudam a recordar-se da toma da medicação? (caso ninguém tenha conhecimento da sua intenção assinalar "não se aplica")

Não se Aplica	Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



2. Em geral, qual é a sua satisfação com o apoio que recebe por parte dos seus amigos e familiares?

Muito Insatisfeito	Insatisfeito	Nem Satisfeito, Nem Insatisfeito	Satisfeito	Muito Satisfeito
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. A seguir vai encontrar várias afirmações sobre relações interpessoais. Pedimos que marque com uma cruz (X) a opção que melhor qualifica a sua forma de pensar.

Por exemplo:
Na primeira afirmação, se você pensa sempre que por vezes se sente só no mundo e sem apoio, deverá assinalar **Concordo Totalmente**, se acha que nunca pensa isso, deverá assinalar **Discordo Totalmente**.

Tendo em conta a última semana	Discorda Plenamente	Discorda	Não Concorda Nem Discorda	Concorda	Concorda Plenamente
Por vezes sinto-me só no mundo e sem apoio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Não saio com amigos tantas vezes quantas eu gostaria	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Os amigos não me procuram tantas vezes quantas eu gostaria	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Quando preciso de desabafar com alguém encontro facilmente amigos com quem o fazer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mesmo nas situações mais embaraçosas, se precisar de apoio de emergência tenho várias pessoas a quem posso recorrer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
As vezes sinto falta de alguém verdadeiramente íntimo que me compreenda e com quem possa desabafar sobre coisas íntimas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sinto falta de actividades sociais que me satisfaçam	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gostava de participar mais em actividades de organizações (p.ex. clubes desportivos, partidos políticos, etc)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Estou satisfeito com a forma como me relaciono com a minha família	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Estou satisfeito com a quantidade de tempo que passo com a minha família	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Estou satisfeito com o que faço em conjunto com a minha família	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Estou satisfeito com a quantidade de amigos que tenho	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Estou satisfeito com a quantidade de tempo que passo com os meus amigos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Estou satisfeito com as actividades e coisas que faço com o meu grupo de amigos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Estou satisfeito com o tipo de amigos que tenho	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



4. Por último, pedimos-lhe que tendo em conta os últimos 7 dias e de uma forma geral, leia cada uma das afirmações abaixo e marque com uma X a opção que mais se aplicou a si. Não leve muito tempo a indicar a sua resposta em cada afirmação.

A afirmação...	Não se Aplicou a mim	Aplicou-se a mim Algumas Vezes	Aplicou-se a mim Muitas Vezes	Aplicou-se a mim Maior Parte das Vezes
...Tive dificuldades em me acalmar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti a minha boca seca	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Não consegui sentir nenhum sentimento positivo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti dificuldades em respirar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Tive dificuldade em tomar iniciativa para fazer coisas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Tive tendência a reagir em demasia em determinadas situações	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti tremores (por ex., nas mãos)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti que estava a utilizar muita energia nervosa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Preocupe-me com situações em que podia entrar em pânico e fazer figura ridícula	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti que não tinha nada a esperar do futuro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Dei por mim a ficar agitado	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti dificuldade em me relaxar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti-me desanimado e melancólico	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Estive intolerante em relação a qualquer coisa que me impedisse de terminar aquilo que estava a fazer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti-me quase a entrar em pânico	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Não fui capaz de ter entusiasmo por nada	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti que não tinha muito valor como pessoa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti que por vezes estava sensível	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti alterações no meu coração sem fazer exercício físico	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti-me assustado sem ter tido uma boa razão para isso	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... Senti que a vida não tinha sentido	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. Até que ponto as perguntas do questionário foram claras para si?

Pouco Claras	Claras	Muito Claras
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Obrigada pela sua participação.



6. Por favor indique "Sim" ou "Não" para cada pergunta, tendo em conta os últimos 30 dias.

	Sim	Não
Consumiu maquiagem durante os últimos 30 dias?	<input type="checkbox"/>	<input type="checkbox"/>
Consumiu anfetaminas (speed) durante os últimos 30 dias?	<input type="checkbox"/>	<input type="checkbox"/>
Consumiu cocaína durante os últimos 30 dias?	<input type="checkbox"/>	<input type="checkbox"/>
Consumiu heroína durante os últimos 30 dias?	<input type="checkbox"/>	<input type="checkbox"/>

7. Por favor marque "Sim" ou "Não"

	Sim	Não
Está actualmente a fazer tratamento com Metadona?	<input type="checkbox"/>	<input type="checkbox"/>

Bloco B

Responda às seguintes perguntas colocando um "X" no quadrado que achar mais apropriado.

Por favor escolha apenas uma opção para cada pergunta

As pessoas têm muitas vezes dificuldade em tomar a sua medicação de acordo com as indicações do médico. Por exemplo, algumas pessoas não conseguem tomar todos os comprimidos receitados no número de vezes indicadas.

1. Nas afirmações que são apresentadas de seguida, qual a que melhor o descreve **actualmente** em relação à toma da medicação anti-retrovírica?

- ☐ Eu não estou a tomar medicação e não estou a pensar começar a tomá-la como indicado
- ☐ Eu não estou a tomar medicação mas estou a pensar começar a tomá-la como indicado
- ☐ Eu não estou a tomar medicação mas vou começar a tomá-la como indicado
- ☐ Eu estou a tomar medicação como indicado.

2.

	Nada/ Nenhum(a)	Pouca(a)	Algum(a)	Bastante	Muito(a)
Como avalia a quantidade de informação que tem sobre a medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Como avalia os benefícios que a medicação lhe pode trazer?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Em quanto é que a sua saúde melhorou desde que começou a medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Qual o grau de esforço que faz para cumprir a toma da sua medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Como avalia a certeza de ser capaz de continuar com a medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Quanto tempo gasta, relacionado com a toma da medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Como avalia a intensidade dos efeitos secundários da medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Qual o grau de dificuldade que sente ao tomar a medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



3.

	Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
Desde que iniciou o tratamento, em alguma ocasião deixou de tomar a sua medicação um dia completo ou mais?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Normalmente, toma a medicação à hora certa?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Quando os resultados das análises são bons, o seu médico costuma utilizá-los para o motivar a continuar o tratamento?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. Quando foi a última vez em que falhou alguma toma da medicação?

- ☐ Durante a última semana
- ☐ Há 1 ou 2 semanas atrás
- ☐ Há 2 ou 4 semanas atrás
- ☐ Há 1 ou 3 meses atrás
- ☐ Há mais de 3 meses
- ☐ Nunca deixou de tomar a medicação

5. As pessoas podem ter dificuldade em tomar todos os medicamentos anti-retrovíricos. Na tabela que se seguinte indique:

- Por favor o nome dos medicamentos anti-retrovíricos que está a tomar neste momento. (Se não se lembrar do nome de algum medicamento, escreva "Não sei" na linha correspondente)
- Para cada medicamento, indique quantas doses não tomou ou tomou fora de horas nos diferentes períodos.

	Nome dos Medicamentos Anti-Retrovíricos	Quantas Doses não tomou ou tomou fora de horas ...				
		Ontem	Há 3 Dias Atrás	Há 4 Dias Atrás	Há 7 Dias Atrás	
1-		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2-		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3-		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4-		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5-		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6-		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Não escrever abaixo desta linha, por favor.

1-

2-

3-

4-

5-

6-



6. Existe algum medicamento que seja mais difícil tomar?

☐ Sim ☐ Não -> Passe para a Pergunta 7

6.b. Se sim, qual(ais) o(s) nome(s)?

6.c. Porque motivo é difícil tomar esse medicamento? (indique por favor a opção mais relevante.)

- ☐ Sabor
- ☐ Horário
- ☐ Receio de efeitos adversos
- ☐ Incompatibilidade com o consumo de alimentos
- ☐ Incompatibilidade com o consumo de álcool/drogas
- ☐ Outro motivo. Qual?

Gostáramos de conhecer melhor os eventuais problemas que tenha tido com a toma da medicação, em diferentes períodos de tempo.

7. Considere agora os **últimos 7 dias**, contando com o dia de ontem.

7.a. Em alguma ocasião esqueceu-se de tomar a sua medicação?

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7.b. Se em alguma ocasião esqueceu-se de tomar a sua medicação, mesmo que raramente, indique o motivo:

7.b. Em alguma ocasião tomou a sua medicação fora de horas?

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7.b. Se em alguma ocasião tomou a sua medicação fora de horas, mesmo que raramente, indique o motivo:

7.c. Nos **últimos 7 dias**, contando com o dia de ontem.

Em alguma ocasião...	Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
... por se ter sentido melhor, deixou de tomar a sua medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... por se ter sentido pior depois de tomar a medicação, deixou de a tomar?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... por se ter sentido triste ou deprimido, deixou de tomar a sua medicação?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. Algumas pessoas esquecem-se de tomar a sua medicação durante os fins-de-semana. No seu caso, falhou alguma da sua medicação anti-retrovírica no último fim-de-semana (último sábado ou domingo)?

☐ Sim
☐ Não

63906

5 / 11



9. Considerando agora os **últimos 4 dias**, indique **quantos dias não tomou a sua medicação?**

Caso tenha tomado apenas uma parte da medicação (por exemplo, apenas 1 comprimido, quando teria de tomar 3 comprimidos) em um ou mais dias, por favor considere esses dias como não tendo tomado a sua medicação.

☐ Nenhum dia ☐ 1 dia ☐ 2 dias ☐ 3 dias ☐ 4 dias

9.a. A maior parte da medicação anti-retrovírica tem que ser tomada segundo um esquema, por ex. "2 vezes por dia" ou "3 vezes por dia" ou "1 de 8 em 8 horas". Considerando os **últimos 4 dias**, com que frequência cumpriu o esquema específico da sua medicação?

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9.b. Alguma da sua medicação anti-retrovírica tem instruções especiais, tais como "tomar com comida" ou "com o estômago vazio" ou "com muitos líquidos"?

☐ Sim ☐ Não

9.c. Se sim, com que frequência seguiu essas instruções especiais nos **últimos 4 dias**?

Nunca	Raramente	Ocasionalmente	Frequentemente	Sempre
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. Em geral, como avalia a sua satisfação com a medicação desde que a começou a tomar?

Muito Insatisfeito	Insatisfeito	Nem Satisfeito, Nem Insatisfeito	Satisfeito	Muito Satisfeito
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

11. Que avaliação faz de si mesmo relativamente à toma da medicação?

Nada Cumpridor	Pouco Cumpridor	Moderadamente Cumpridor	Bastante Cumpridor	Muito Cumpridor
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12. Alguma vez recebeu algum tipo de apoio para a toma de medicação anti-retrovírica?

☐ Sim ☐ Não

12.a. Se sim, indique por favor de quem recebeu apoio? (Pode assinalar mais que uma opção).

☐ Familiar ☐ Psicólogo/a na consulta ☐ Farmacêutico ☐ Enfermeiro(a)

☐ Amigos ☐ Consulta de Adesão ☐ Médico(a) ☐ Outro Qual?

12.b. Actualmente está a receber algum tipo de apoio para a toma da medicação?

☐ Sim ☐ Não

12.c. Se sim, indique por favor de quem recebeu apoio? (Pode assinalar mais que uma opção).

☐ Familiar ☐ Psicólogo/a na consulta ☐ Farmacêutico ☐ Enfermeiro(a)

☐ Amigos ☐ Consulta de Adesão ☐ Médico(a) ☐ Outro Qual?

63906

6 / 11



Gostariamos de saber a sua opinião acerca da medicação anti-retrovírica

Apresentamos a seguir afirmações que outras pessoas fizeram sobre a sua medicação anti-retrovírica. Indique, por favor, até que ponto concorda ou discorda delas assinalando com uma X resposta que mais se adequa ao seu caso.

13.

	Discorda Plenamente	Discorda	Não Concorda Nem Discorda	Concorda	Concorda Plenamente
Atualmente, a minha saúde depende da medicação anti-retrovírica que me foi prescrita	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Preocupa-me ler que tomar esta medicação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha vida seria impossível sem esta medicação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sem esta medicação eu ficaria muito doente	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Por vezes, preocupo-me com os efeitos a longo prazo desta medicação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esta medicação é um mistério para mim	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
No futuro, a minha saúde vai depender desta medicação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esta medicação perturba a minha vida	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Por vezes, preocupo-me poder ficar demasiado dependente desta medicação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esta medicação protege-me de ficar pior	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esta medicação provoca-me efeitos secundários desagradáveis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Para mim, é difícil tomar a medicação a horas todos os dias	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Usar esta medicação é embaraçoso	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Não tomar esta medicação durante um dia não tem importância a longo prazo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esta medicação aumenta a esperança no futuro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Conciliar a toma desta medicação com as refeições não constitui problema	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Não é provável que eu tenha um efeito secundário desagradável causado por esta medicação durante o próximo mês	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tomar esta medicação está a ser muito pior do que esperava	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Recebi informação suficiente sobre a medicação anti-retrovírica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esta medicação mantém-me vivo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O gosto da medicação anti-retrovírica faz-me sentir mal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



14.

	Discorda Plenamente	Discorda	Não Concorda Nem Discorda	Concorda	Concorda Plenamente
Eu sei como é suposto tomar cada um dos meus medicamentos anti-retrovíricos (por exemplo: tomar ou não com alimentos ou outros medicamentos)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu sei o que fazer se falhar a minha toma de medicação anti-retrovírica (por exemplo: tomar ou não tomar mais tarde)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Deixar de tomar a medicação anti-retrovírica de vez em quando provavelmente não fará mal à minha saúde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu sei quais os possíveis efeitos secundários da minha medicação anti-retrovírica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Desde que eu me sinta saudável, não há problema em esquecer de tomar a medicação anti-retrovírica de vez em quando	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu percebo como funciona a minha medicação anti-retrovírica no meu corpo, para combater o VIH	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Se eu não tomar a minha medicação anti-retrovírica como prescrita, este tipo de medicação pode vir a não funcionar em mim, no futuro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu acredito que, se tomar a minha medicação anti-retrovírica como foi prescrita, eu vou viver mais tempo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu sei como a minha medicação anti-retrovírica interage com álcool e com drogas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu fico preocupado que as outras pessoas saibam que eu tenho VIH, por me verem a tomar a minha medicação anti-retrovírica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu fico frustrado com a toma da medicação anti-retrovírica porque tenho de planejar a minha vida em função dela	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu não gosto de tomar a medicação anti-retrovírica porque ela me lembra que eu tenho VIH	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu sinto que o meu médico considera o que eu preciso quando me diz quais os medicamentos anti-retrovíricos a tomar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A maioria das pessoas que são importantes para mim e que sabem que tenho VIH, ajuda-me a tomar a minha medicação anti-retrovírica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
O meu médico não me dá o apoio suficiente para a toma da medicação anti-retrovírica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sinto-me frustrado quando penso que terei de tomar esta medicação anti-retrovírica todos os dias, no resto da minha vida.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Eu estou preocupado que esta medicação anti-retrovírica que me foi prescrita me faça mal à saúde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Perturba-me que esta medicação anti-retrovírica prescrita possa afectar o meu aspecto	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Perturba-me que esta medicação anti-retrovírica prescrita possa causar efeitos secundários	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Existem momentos quando bebo álcool ou drogas em que é difícil tomar a medicação anti-retrovírica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



17. Pedimos-lhe que leia com atenção as seguintes afirmações e que assinale com um X a hipótese que melhor descreve em que medida concorda ou discorda em relação a cada afirmação

	Discorda Por completo	Discorda	Não Concorda Nem Discorda	Concorda	Concorda Por completo
Acredito que a minha Equipa de Saúde está bem informada em relação à infecção VIH/SIDA.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde esforça-se para me tratar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acredito que a minha Equipa de Saúde está motivada em ajudar-me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde preocupa-se com a minha saúde em geral	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acredito que a minha Equipa de Saúde sabe bastante sobre medicamentos para o VIH	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acredito que recebo os melhores cuidados de saúde disponíveis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde é preguiçosa.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde está bem informada em relação a novos tratamentos para o VIH.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acredito que a minha Equipa de Saúde se preocupa comigo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde dá-me apoio/suporte.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde encoraja-me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde é prestável.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde faz-me sentir à vontade.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde está comigo o tempo suficiente.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde é sensível em relação ao que sinto.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde acha-me uma má pessoa porque tenho VIH.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde dá importância à minha opinião.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acredito que a minha Equipa de Saúde me acha estúpido.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A minha Equipa de Saúde faz juízos sobre mim.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

15. Até que ponto é difícil ou fácil para si...

Muito Difícil	Difícil	Às Vezes É Fácil ou Difícil	Fácil	Muito Fácil
...estar informado sobre a medicação anti-retrovírica?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...ter a ajuda necessária dos outros para tomar a medicação anti-retrovírica (por exemplo: amigos, família, médico, farmacêutico, ...)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...ir levantar a sua medicação anti-retrovírica sem a deixar acabar?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...tomar a medicação anti-retrovírica quando está envolvido no que está a fazer?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...lidar com os efeitos secundários da toma a medicação anti-retrovírica?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...lembrar-se de tomar a medicação anti-retrovírica?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...tomar a medicação anti-retrovírica quando os comprimidos são difíceis de engolir, sabem mal ou o fazem sentir-se mal do estômago?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...fazer com que a medicação anti-retrovírica faça parte da sua rotina diária?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...tomar a medicação anti-retrovírica quando existe uma mudança na sua rotina diária (por exemplo, quando viaja ou quando vai sair com os amigos...)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...tomar a medicação anti-retrovírica quando não se sente bem emocionalmente (por exemplo, quando está triste, zangado, nervoso ou stressado)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...tomar a medicação anti-retrovírica quando se sente bem fisicamente e não tem nenhum sintoma da infecção por VIH?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...tomar a medicação anti-retrovírica quando não se sente bem fisicamente?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...falar com o seu médico sobre a medicação anti-retrovírica?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
...falar com o seu médico sobre a sua vida sexual, consumo de álcool ou consumo de drogas?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... confiar no seu médico?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

16. Da seguinte lista, assinale os profissionais de saúde que fazem parte da sua Equipa de Saúde que o acompanha neste serviço, quanto à infecção por VIH.

☐ Médico

☐ Psicólogo

☐ Técnico de Serviço Social

☐ Enfermeiro

☐ Outro. Qual

☐ Nutricionista

☐ Não sabe / Não responde

☐ Farmacêutico

Estudo sobre utilização de medicação anti-retrovírica

18. Como avalia a sua satisfação com o médico que o acompanha para o VIH em relação ...

	Muito Insatisfeito	Insatisfeito	Nem Satisfeito, Nem Insatisfeito	Satisfeito	Muito Satisfeito
... à informação dada sobre como tomar correctamente a sua medicação anti-retrovírica?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... à compreensão dos seus problemas com a toma da medicação anti-retrovírica?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... à ajuda dada para resolver os seus problemas com a toma correcta da medicação anti-retrovírica?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... à forma como explica os resultados dos testes laboratoriais?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... à explicação sobre os benefícios e os riscos com o tratamento?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... à informação dada sobre o que fazer se acontecerem determinados sintomas ou problemas?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... ao apoio que lhe demonstra?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... à compreensão das suas dúvidas e preocupações com a saúde?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

19. Até que ponto está satisfeito com a relação que tem com o médico que o acompanha na infecção VIH/SIDA?

Muito Insatisfeito	Insatisfeito	Nem Satisfeito, Nem Insatisfeito	Satisfeito	Muito Satisfeito
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

20. Recomendaria o médico que o acompanha para a infecção VIH a outras pessoas?

Definitivamente Não	Provavelmente Não	Não tenho a certeza	Provavelmente Sim	Definitivamente Sim
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

21. Chegámos ao final deste questionário dando por concluída a sua participação.

Para melhor planeamento de estudos futuros, até que ponto as perguntas foram claras para si?

Nada	Pouco	Alguns	Bastante	Muito
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

22. Estaria disponível para continuar a participar neste estudo? ☐ Sim ☐ Não22.a. Se sim, por mais quantos meses? meses

Muito Obrigada pela sua participação no estudo!

11 / 11



Instituto de Medicina Preventiva
Faculdade de Medicina de Lisboa



ESTUDO DA UTILIZAÇÃO DA MEDICAÇÃO ANTI-RETROVÍRICA

CONSENTIMENTO INFORMADO

Este documento pretende dar-lhe a **informação** essencial de que depende o seu **consentimento** para participar neste estudo de investigação e é um requisito necessário para essa **participação**.

Pede-se que o **leia**, coloque as suas **dúvidas** ao seu médico assistente e, se aceitar participar, o **assine**.

Leve o tempo que entender necessário para examiná-lo.

Desde já, fica o nosso obrigado!

Natureza, objectivos e procedimentos do estudo

O nosso **objectivo** é **estudar os medicamentos usados na infecção VIH/SIDA**, recolhendo **informação** sobre o tipo de medicamento, a forma como é tomado e os seus efeitos.

Para termos essa **informação**,

1. pedimos que preencha um **Questionário**, a seguir a esta consulta. O questionário tem perguntas sobre si e **sobre a sua medicação anti-retrovírica**, sem que hajam respostas certas ou erradas. O questionário é **confidencial**. Terá o apoio de uma psicóloga, se necessitar de esclarecer qualquer questão.

2. será consultado o seu **Processo Clínico e o registo da dispensa de medicação da farmácia hospitalar** (pelo seu médico assistente ou por um profissional de saúde autorizado por ele e sujeito a segredo profissional)

Para conhecer melhor as alterações na forma como a medicação é tomada, quais os efeitos que aparecem e como isso está relacionado com o controlo da doença, **precisamos de o acompanhar durante 6 meses**.

Assim, será pedido que preencha um **novo questionário quando vier à próxima consulta agendada**. Se por qualquer motivo tiver de faltar à consulta, iremos tentar contactá-lo/a por telefone para que possa responder ao questionário. Estes questionários e a consulta do registo e do processo clínico representam a totalidade da sua participação neste estudo e **nada mais lhe será pedido**.

As suas respostas serão tratadas, publicadas e apresentadas sempre de forma agrupada com as respostas dos outros participantes e **nunca de forma individual ou que o/a possa identificar**.

Os dados recolhidos serão **acedidos apenas** pelos investigadores e guardados durante o período de realização do estudo no Instituto de Medicina Preventiva, da Faculdade de Medicina de Lisboa, sendo posteriormente destruídos. Os investigadores assumem a **responsabilidade pela confidencialidade** de quaisquer dados recolhidos.

A sua participação neste estudo **não envolve** riscos ou custos para si e não estão previstas gratificações. Este estudo **não tem** objectivos comerciais e será importante para médicos, doentes e decisores políticos.

Dado o carácter **voluntário** da sua participação, pode desvincular-se a qualquer momento do estudo. A recusa inicial ou o abandono subsequente não acarretam quaisquer penalizações ou perda de direitos.

Declaro ter sido devidamente informado/a sobre o estudo e aceitar participar

Lisboa, ____/____/2011

Assinatura do/a Participante: _____

Assinatura do Investigador Responsável: _____

Conserve esta cópia do documento, ficando a outra cópia na posse do seu médico assistente.

Se tiver alguma dúvida ou vier a precisar de algum esclarecimento, poderá contactar **Dr.ª Milene Fernandes** (coordenadora do estudo)
Instituto de Medicina Preventiva - Faculdade de Medicina de Lisboa - Tlf 21 799 94 22, ext. 47090 - E-mail: mci@fm.ul.pt

O seu código de participação no estudo é: | ____ | ____ | ____ | ____ | ____ | ____ | ____ | ____ | ____ | ____ |

Para fins do estudo, pedimos-lhe que nos indique o seu contacto telefónico _____

63906



FORMULÁRIO DE ELEGIBILIDADE

Nome: _____

NID:

Sexo ☐ Masculino ☐ Feminino

Data Nascimento: DD MM AAAA

O doente cumpre todos os critérios abaixo indicados? ☐ Sim ☐ Não

Se não cumpre todos os critérios assinala com, um x os critérios em falta

1. Seropositivo para VIH 1

☐

2. Idade superior ou igual a 18 anos

☐

3. Seguidos neste hospital e neste serviço

☐

4. Em terapêutica anti-retrovírica ou com indicação para iniciá-la

☐

5. Com capacidade de prestar consentimento informado

☐

6. Não está internado

☐

7. Não é participante num ensaio clínico

☐

8. Não está em nenhum programa de toma sob observação directa (para a toma de anti-retrovirais)

☐

9. Não veio de nem é seguido noutra hospital (para a infecção VIH)

☐

10. Não depende de terceiros para a toma da medicação (família, prisão, instituição social)

☐

Recusou participar

☐ Não ☐ Sim

Motivo da recusa

3.a. se o doente for elegível

- preencher os campos NOME e NID (não entram na base de dados), Sexo e Data de Nascimento

- deixe ficar este formulário no processo clínico do doente e entregue-o juntamente com o consentimento informado a entrevistadora, após a consulta.

3.b. se o doente não for elegível

- preencher apenas os campos Sexo e Data de Nascimento – não preencher os campos NOME e NID

- entregue o formulário durante a próxima semana às entrevistadoras

4. Se o doente for elegível mas recusar participar, assinale no campo RECUSA e entregue o formulário às entrevistadoras.

Para o esclarecimento de qualquer questão, poderá contactar uma das entrevistadoras ou a coordenadora do estudo (Mylene Fernandes 965751328)

Obrigada pela colaboração!

Não preencher os campos abaixo desta linha, por favor

ID Participante

ID Questionário

21376

